



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 113138

TO: Rita Mitra
Location:
Art Unit: 1653
Friday, January 30, 2004

Case Serial Number: 10/032950

From: Toby Port
Location: Biotech-Chem Library
CM1-6A04
Phone: 308-3534

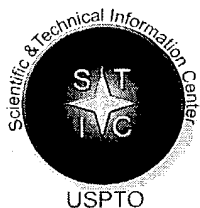
toby.port@uspto.gov

Search Notes

Dear Examiner Mitra,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



=> file caplus; d que l13; d que l16; d que l18; d que l21; d que l24; d que l25; d que l26

FILE 'CAPLUS' ENTERED AT 17:33:05 ON 30 JAN 2004

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

COPYRIGHT (C) 2004 AMERICAN CHEMICAL SOCIETY (ACS)

Copyright of the articles to which records in this database refer is held by the publishers listed in the PUBLISHER (PB) field (available for records published or updated in Chemical Abstracts after December 26, 1996), unless otherwise indicated in the original publications. The CA Lexicon is the copyrighted intellectual property of the American Chemical Society and is provided to assist you in searching databases on STN. Any dissemination, distribution, copying, or storing of this information, without the prior written consent of CAS, is strictly prohibited.

FILE COVERS 1907 - 30 Jan 2004 VOL 140 ISS 6

FILE LAST UPDATED: 29 Jan 2004 (20040129/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

L3	21584	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	STAPHYLOCOCCUS AUREUS/CT
L4	700510	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	PROTEINS/CT
L5	117619	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	PEPTIDES/CW
L6	72487	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	ANTIBACTERIAL AGENTS+OLD/CT
L7	80387	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	(L4 OR L5) (L) (PREP OR SPN OR THU)/RL
L12	15522	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	(L4 OR L5) (L) CYCL?
L13	13	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	L3 AND L7 AND L12 AND L6

L3	21584	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	STAPHYLOCOCCUS AUREUS/CT
L4	700510	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	PROTEINS/CT
L5	117619	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	PEPTIDES/CW
L6	72487	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	ANTIBACTERIAL AGENTS+OLD/CT
L7	80387	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	(L4 OR L5) (L) (PREP OR SPN OR THU)/RL
L10	94092	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	LACTAM OR LACTONE
L12	15522	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	(L4 OR L5) (L) CYCL?
L13	13	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	L3 AND L7 AND L12 AND L6
L14	14	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	L3 AND (L7 OR L12) AND L10
L15	13	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	L14 NOT L13
L16	1	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	L15 AND DIRECTED/TI

L3	21584	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	STAPHYLOCOCCUS AUREUS/CT
L4	700510	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	PROTEINS/CT
L5	117619	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	PEPTIDES/CW
L11	2682	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	AGR
L12	15522	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	(L4 OR L5) (L) CYCL?
L18	3	SEA	FILE=CAPLUS	ABB=ON	PLU=ON	L3 AND L11 AND L12

```

L3      21584 SEA FILE=CAPLUS ABB=ON  PLU=ON  STAPHYLOCOCCUS AUREUS/CT
L4      700510 SEA FILE=CAPLUS ABB=ON  PLU=ON  PROTEINS/CT
L5      117619 SEA FILE=CAPLUS ABB=ON  PLU=ON  PEPTIDES/CW
L10     94092 SEA FILE=CAPLUS ABB=ON  PLU=ON  LACTAM OR LACTONE
L11     2682 SEA FILE=CAPLUS ABB=ON  PLU=ON  AGR
L12     15522 SEA FILE=CAPLUS ABB=ON  PLU=ON  (L4 OR L5) (L) CYCL?
L18      3 SEA FILE=CAPLUS ABB=ON  PLU=ON  L3 AND L11 AND L12
L20     2 SEA FILE=CAPLUS ABB=ON  PLU=ON  L3 AND L11 AND ((L4 OR L5) OR
CYCL?) AND L10
L21     2 SEA FILE=CAPLUS ABB=ON  PLU=ON  L20 NOT L18

```

```

L3      21584 SEA FILE=CAPLUS ABB=ON  PLU=ON  STAPHYLOCOCCUS AUREUS/CT
L4      700510 SEA FILE=CAPLUS ABB=ON  PLU=ON  PROTEINS/CT
L5      117619 SEA FILE=CAPLUS ABB=ON  PLU=ON  PEPTIDES/CW
L12     15522 SEA FILE=CAPLUS ABB=ON  PLU=ON  (L4 OR L5) (L) CYCL?
L22      69 SEA FILE=CAPLUS ABB=ON  PLU=ON  ACCESSORY GENE REGULAT?
L24      0 SEA FILE=CAPLUS ABB=ON  PLU=ON  L3 AND L22 AND L12

```

```

L3      21584 SEA FILE=CAPLUS ABB=ON  PLU=ON  STAPHYLOCOCCUS AUREUS/CT
L6      72487 SEA FILE=CAPLUS ABB=ON  PLU=ON  ANTIBACTERIAL AGENTS+OLD/CT
L22      69 SEA FILE=CAPLUS ABB=ON  PLU=ON  ACCESSORY GENE REGULAT?
L25      0 SEA FILE=CAPLUS ABB=ON  PLU=ON  L3 AND L22 AND L6

```

```

L3      21584 SEA FILE=CAPLUS ABB=ON  PLU=ON  STAPHYLOCOCCUS AUREUS/CT
L10     94092 SEA FILE=CAPLUS ABB=ON  PLU=ON  LACTAM OR LACTONE
L22      69 SEA FILE=CAPLUS ABB=ON  PLU=ON  ACCESSORY GENE REGULAT?
L26      0 SEA FILE=CAPLUS ABB=ON  PLU=ON  L3 AND L22 AND L10

```

```

=> s l13 or l16 or l18 or l21
L72     19 L13 OR L16 OR L18 OR L21

```

```

=> file medline; d que l35; d que l37; d que l38
FILE 'MEDLINE' ENTERED AT 17:34:18 ON 30 JAN 2004

```

FILE LAST UPDATED: 29 JAN 2004 (20040129/UP). FILE COVERS 1958 TO DATE.

On December 14, 2003, the 2004 MeSH terms were loaded. See HELP RLOAD for details.

MEDLINE thesauri in the /CN, /CT, and /MN fields incorporate the MeSH 2004 vocabulary. See <http://www.nlm.nih.gov/mesh/> and http://www.nih.gov/pubs/yechebull/nd03/nd03_mesh.html for a description on changes.

This file contains CAS Registry Numbers for easy and accurate substance identification.

```

L28     24379 SEA FILE=MEDLINE ABB=ON  PLU=ON  STAPHYLOCOCCUS AUREUS/CT
L29     497 SEA FILE=MEDLINE ABB=ON  PLU=ON  AGR OR ACCESSORY GENE
REGULAT?
L30     107060 SEA FILE=MEDLINE ABB=ON  PLU=ON  ANTI-BACTERIAL AGENTS/CT OR
ANTI-INFECTIVE AGENTS/CT
L35      5 SEA FILE=MEDLINE ABB=ON  PLU=ON  L28 AND L29 AND L30

```


L28 24379 SEA FILE=MEDLINE ABB=ON PLU=ON STAPHYLOCOCCUS AUREUS/CT
 L29 497 SEA FILE=MEDLINE ABB=ON PLU=ON AGR OR ACCESSORY GENE
 REGULAT?
 L31 354100 SEA FILE=MEDLINE ABB=ON PLU=ON CYCLIC (3A) PROTEIN OR
 PEPTIDE
 L33 56475 SEA FILE=MEDLINE ABB=ON PLU=ON LACTONES+NT/CT
 L37 2 SEA FILE=MEDLINE ABB=ON PLU=ON L28 AND L29 AND L31 AND L33

L28 24379 SEA FILE=MEDLINE ABB=ON PLU=ON STAPHYLOCOCCUS AUREUS/CT
 L29 497 SEA FILE=MEDLINE ABB=ON PLU=ON AGR OR ACCESSORY GENE
 REGULAT?
 L31 354100 SEA FILE=MEDLINE ABB=ON PLU=ON CYCLIC (3A) PROTEIN OR
 PEPTIDE
 L34 72703 SEA FILE=MEDLINE ABB=ON PLU=ON LACTAMS+NT/CT
 L38 3 SEA FILE=MEDLINE ABB=ON PLU=ON L28 AND L29 AND L31 AND L34

=> 's l35 or l37 or l38
 L73 9 L35 OR L37 OR L38

=> file embase; d que l48; d que l49; d que l50
 FILE 'EMBASE' ENTERED AT 17:34:59 ON 30 JAN 2004
 COPYRIGHT (C) 2004 Elsevier Inc. All rights reserved.

FILE COVERS 1974 TO 29 Jan 2004 (20040129/ED)

EMBASE has been reloaded. Enter HELP RLOAD for details.

This file contains CAS Registry Numbers for easy and accurate
 substance identification.

L39 35935 SEA FILE=EMBASE ABB=ON PLU=ON "STAPHYLOCOCCUS AUREUS"+ALL/CT
 L40 838278 SEA FILE=EMBASE ABB=ON PLU=ON ANTIINFECTIVE AGENT+NT/CT
 L41 363 SEA FILE=EMBASE ABB=ON PLU=ON AGR OR ACCESSORY GENE REGULAT?
 L42 30871 SEA FILE=EMBASE ABB=ON PLU=ON (PROTEIN OR PEPTIDE) (3A)
 CYCL?
 L48 0 SEA FILE=EMBASE ABB=ON PLU=ON L39 AND L41 AND L40 AND L42

L39 35935 SEA FILE=EMBASE ABB=ON PLU=ON "STAPHYLOCOCCUS AUREUS"+ALL/CT
 L40 838278 SEA FILE=EMBASE ABB=ON PLU=ON ANTIINFECTIVE AGENT+NT/CT
 L41 363 SEA FILE=EMBASE ABB=ON PLU=ON AGR OR ACCESSORY GENE REGULAT?
 L44 4193 SEA FILE=EMBASE ABB=ON PLU=ON LACTAM+NT/CT
 L45 7912 SEA FILE=EMBASE ABB=ON PLU=ON LACTONE+NT/CT
 L49 0 SEA FILE=EMBASE ABB=ON PLU=ON L39 AND L41 AND L40 AND (L44
 OR L45)

L39 35935 SEA FILE=EMBASE ABB=ON PLU=ON "STAPHYLOCOCCUS AUREUS"+ALL/CT

L40 838278 SEA FILE=EMBASE ABB=ON PLU=ON ANTIINFECTIVE AGENT+NT/CT
 L41 363 SEA FILE=EMBASE ABB=ON PLU=ON AGR OR ACCESSORY GENE REGULAT?
 L50 0 SEA FILE=EMBASE ABB=ON PLU=ON L39 AND L41 AND L40 AND CYCL?

=> file wpids; d que 157; d que 158; d que 160
 FILE 'WPIDS' ENTERED AT 17:35:16 ON 30 JAN 2004
 COPYRIGHT (C) 2004 THOMSON DERWENT

FILE LAST UPDATED: 28 JAN 2004 <20040128/UP>
 MOST RECENT DERWENT UPDATE: 200407 <200407/DW>
 DERWENT WORLD PATENTS INDEX SUBSCRIBER FILE, COVERS 1963 TO DATE

>>> NEW WEEKLY SDI FREQUENCY AVAILABLE --> see NEWS <<<

>>> PATENT IMAGES AVAILABLE FOR PRINT AND DISPLAY <<<

>>> FOR A COPY OF THE DERWENT WORLD PATENTS INDEX STN USER GUIDE,
 PLEASE VISIT:
http://www.stn-international.de/training_center/patents/stn_guide.pdf <<<

>>> FOR DETAILS OF THE PATENTS COVERED IN CURRENT UPDATES, SEE
<http://thomsonderwent.com/coverage/latestupdates/> <<<

>>> FOR INFORMATION ON ALL DERWENT WORLD PATENTS INDEX USER
 GUIDES, PLEASE VISIT:
<http://thomsonderwent.com/support/userguides/> <<<

>>> ADDITIONAL POLYMER INDEXING CODES WILL BE IMPLEMENTED FROM
 DERWENT UPDATE 200403.
 THE TIME RANGE CODE WILL ALSO CHANGE FROM 018 TO 2004.
 SDIS USING THE TIME RANGE CODE WILL NEED TO BE UPDATED.
 FOR FURTHER DETAILS: <http://thomsonderwent.com/chem/polymers/> <<<

L51 3758 SEA FILE=WPIDS ABB=ON PLU=ON STAPHYLOCOCCUS AUREUS
 L52 1246 SEA FILE=WPIDS ABB=ON PLU=ON S AUREUS
 L53 92 SEA FILE=WPIDS ABB=ON PLU=ON AGR OR ACCESSORY GENE REGULAT?
 L54 48087 SEA FILE=WPIDS ABB=ON PLU=ON ANTIBACTER? OR ANTIINFECT? OR
 BACTERICID? OR ANTI (W) (BACTER? OR INFECT?)
 L57 7 SEA FILE=WPIDS ABB=ON PLU=ON (L51 OR L52) AND L53 AND L54

L51 3758 SEA FILE=WPIDS ABB=ON PLU=ON STAPHYLOCOCCUS AUREUS
 L52 1246 SEA FILE=WPIDS ABB=ON PLU=ON S AUREUS
 L53 92 SEA FILE=WPIDS ABB=ON PLU=ON AGR OR ACCESSORY GENE REGULAT?
 L55 19291 SEA FILE=WPIDS ABB=ON PLU=ON LACTAM OR LACTONE
 L58 2 SEA FILE=WPIDS ABB=ON PLU=ON (L51 OR L52) AND L53 AND L55

L51 3758 SEA FILE=WPIDS ABB=ON PLU=ON STAPHYLOCOCCUS AUREUS
 L52 1246 SEA FILE=WPIDS ABB=ON PLU=ON S AUREUS
 L53 92 SEA FILE=WPIDS ABB=ON PLU=ON AGR OR ACCESSORY GENE REGULAT?
 L56 2040 SEA FILE=WPIDS ABB=ON PLU=ON (PEPTIDE OR PROTEIN) (3A) CYCL?

L60 4 SEA FILE=WPIDS ABB=ON PLU=ON (L51 OR L52) AND L53 AND L56

=> s 157 or 158 or 160

L74 7 L57 OR L58 OR L60

=> file biosis; d que 168; d que 171

FILE 'BIOSIS' ENTERED AT 17:35:48 ON 30 JAN 2004

COPYRIGHT (C) 2004 BIOLOGICAL ABSTRACTS INC.(R)

FILE COVERS 1969 TO DATE.

CAS REGISTRY NUMBERS AND CHEMICAL NAMES (CNs) PRESENT
FROM JANUARY 1969 TO DATE.

RECORDS LAST ADDED: 28 January 2004 (20040128/ED)

FILE RELOADED: 19 October 2003.

L62 60064 SEA FILE=BIOSIS ABB=ON PLU=ON (STAPHYLOCOCCUS OR STAPH OR S)
(W) AUREUS

L63 604 SEA FILE=BIOSIS ABB=ON PLU=ON AGR OR ACCESSORY GENE REGULAT?

L64 204285 SEA FILE=BIOSIS ABB=ON PLU=ON ANTIBACTER? OR ANTIINFECT? OR
ANTI (W) (BACTER? OR INFECT?) OR BACTERICID?

L67 26 SEA FILE=BIOSIS ABB=ON PLU=ON L62 AND L63 AND L64

L68 1 SEA FILE=BIOSIS ABB=ON PLU=ON L67 AND ENGINEERED/TI

L62 60064 SEA FILE=BIOSIS ABB=ON PLU=ON (STAPHYLOCOCCUS OR STAPH OR S)
(W) AUREUS

L63 604 SEA FILE=BIOSIS ABB=ON PLU=ON AGR OR ACCESSORY GENE REGULAT?

L66 29093 SEA FILE=BIOSIS ABB=ON PLU=ON LACTONE OR LACTAM

L70 5 SEA FILE=BIOSIS ABB=ON PLU=ON L62 AND L63 AND L66

L71 1 SEA FILE=BIOSIS ABB=ON PLU=ON L70 AND EVOLUTION/TI

=> s 168 or 171

L75 2 L68 OR L71

=> dup rem 173 172 175 174

FILE 'MEDLINE' ENTERED AT 17:36:22 ON 30 JAN 2004

FILE 'CAPLUS' ENTERED AT 17:36:22 ON 30 JAN 2004

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

COPYRIGHT (C) 2004 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'BIOSIS' ENTERED AT 17:36:22 ON 30 JAN 2004

COPYRIGHT (C) 2004 BIOLOGICAL ABSTRACTS INC.(R)

FILE 'WPIDS' ENTERED AT 17:36:22 ON 30 JAN 2004

COPYRIGHT (C) 2004 THOMSON DERWENT

PROCESSING COMPLETED FOR L73

PROCESSING COMPLETED FOR L72

PROCESSING COMPLETED FOR L75

PROCESSING COMPLETED FOR L74

L76 34 DUP REM L73 L72 L75 L74 (3 DUPLICATES REMOVED)

ANSWERS '1-9' FROM FILE MEDLINE

ANSWERS '10-27' FROM FILE CAPLUS
ANSWER '28' FROM FILE BIOSIS
ANSWERS '29-34' FROM FILE WPIDS

=> d ibib ab 176 1-34

L76 ANSWER 1 OF 34 MEDLINE on STN DUPLICATE 2
ACCESSION NUMBER: 2001444137 MEDLINE
DOCUMENT NUMBER: 21382775 PubMed ID: 11489134
TITLE: Structure, activity and evolution of the group I
thiolactone **peptide** quorum-sensing system of
Staphylococcus aureus.
AUTHOR: MDowell P; Affas Z; Reynolds C; Holden M T; Wood S J; Saint
S; Cockayne A; Hill P J; Dodd C E; Bycroft B W; Chan W C;
Williams P
CORPORATE SOURCE: Institute of Infections and Immunity, School of
Pharmaceutical Sciences, and School of BioSciences,
University of Nottingham, Nottingham NG7 2RD, UK.
SOURCE: MOLECULAR MICROBIOLOGY, (2001 Jul) 41 (2) 503-12.
Journal code: 8712028. ISSN: 0950-382X.
PUB. COUNTRY: England: United Kingdom
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
OTHER SOURCE: GENBANK-AF001782; GENBANK-AF001783
ENTRY MONTH: 200204
ENTRY DATE: Entered STN: 20010813
Last Updated on STN: 20020412
Entered Medline: 20020411

AB In Staphylococcus aureus, the **agr** locus is responsible for
controlling virulence gene expression via quorum sensing. As the blockade
of quorum sensing offers a novel strategy for attenuating infection, we
sought to gain novel insights into the structure, activity and turnover of
the secreted staphylococcal autoinducing **peptide** (AIP) signal
molecules. A series of analogues (including the L-alanine and D-amino
acid scanned **peptides**) was synthesized to determine the
functionally critical residues within the S. aureus group I AIP. As a
consequence, we established that (i) the group I AIP is inactivated in
culture supernatants by the formation of the corresponding methionyl
sulphoxide; and (ii) the group I AIP lactam analogue retains the capacity
to activate **agr**, suggesting that covalent modification of the
AgrC receptor is not a necessary prerequisite for **agr**
activation. Although each of the D-amino acid scanned AIP analogues
retained activity, replacement of the endocyclic amino acid residue
(aspartate) located C-terminally to the central cysteine with alanine
converted the group I AIP from an activator to a potent inhibitor. The
screening of clinical S. aureus isolates for novel AIP groups revealed a
variant that differed from the group I AIP by a single amino acid residue
(aspartate to tyrosine) in the same position defined as critical by
alanine scanning. Although this AIP inhibits group I S. aureus strains,
the producer strains possess a functional **agr** locus dependent on
the endogenous **peptide** and, as such, constitute a fourth S.
aureus AIP pheromone group (group IV). The addition of exogenous
synthetic AIPs to S. aureus inhibited the production of toxic shock
syndrome toxin (TSST-1) and enterotoxin C3, confirming the potential of
quorum-sensing blockade as a therapeutic strategy.

L76 ANSWER 2 OF 34 MEDLINE on STN DUPLICATE 3
ACCESSION NUMBER: 1999145530 MEDLINE
DOCUMENT NUMBER: 99145530 PubMed ID: 9990004
TITLE: Structure-activity analysis of synthetic autoinducing

thiolactone **peptides** from *Staphylococcus aureus* responsible for virulence.

AUTHOR: Mayville P; Ji G; Beavis R; Yang H; Goger M; Novick R P; Muir T W

CORPORATE SOURCE: Laboratory of Synthetic Protein Chemistry, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA.

CONTRACT NUMBER: A142783 (NIAID)
AI30138-09 (NIGMS)
GM55843-01

SOURCE: PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, (1999 Feb 16) 96 (4) 1218-23.
Journal code: 7505876. ISSN: 0027-8424.

PUB. COUNTRY: United States

DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)

LANGUAGE: English

FILE SEGMENT: Priority Journals

ENTRY MONTH: 199903

ENTRY DATE: Entered STN: 19990402
Last Updated on STN: 19990402
Entered Medline: 19990325

AB The synthesis of virulence factors and other extracellular proteins responsible for pathogenicity in *Staphylococcus aureus* is under the control of the **agr** locus. A secreted **agr**-encoded **peptide**, AgrD, processed from the AgrD gene product, is known to be an effector of self-strain activation and cross-strain inhibition of the **agr** response. Biochemical analysis of AgrD **peptides** isolated from culture supernatants has suggested that they contain an unusual thiol ester-linked cyclic structure. In the present work, chemical synthesis is used to confirm that the mature AgrD **peptides** contain a thiolactone structure and that this feature is absolutely necessary for full biological activity. The AgrD synthetic thiolactone **peptides** exhibited biological activity in vivo in a mouse protection test. Structure-activity studies have allowed key aspects of the **peptide** structure involved in the differential activation and inhibition functions to be identified. Accordingly, we propose a model for activation and inhibition of the **agr** response in which the former, but not the latter, involves specific acylation of the **agr** transmembrane receptor, AgrC.

L76 ANSWER 3 OF 34 MEDLINE on STN

ACCESSION NUMBER: 2003145749 MEDLINE

DOCUMENT NUMBER: 22546994 PubMed ID: 12660939

TITLE: *Staphylococcus aureus* accessory gene regulator (**agr**) group II: is there a relationship to the development of intermediate-level glycopeptide resistance?.

AUTHOR: Sakoulas George; Eliopoulos George M; Moellering Robert C Jr; Novick Richard P; Venkataraman Lata; Wennersten Christine; DeGirolami Paola C; Schwaber Mitchell J; Gold Howard S

CORPORATE SOURCE: Department of Medicine, Division of Laboratory and Transfusion Medicine and Harvard Medical School, Boston, Massachusetts, USA.. gsakoulas@crystalrunhealthcare.com

SOURCE: JOURNAL OF INFECTIOUS DISEASES, (2003 Mar 15) 187 (6) 929-38.
Journal code: 0413675. ISSN: 0022-1899.

PUB. COUNTRY: United States

DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)

LANGUAGE: English

FILE SEGMENT: Abridged Index Medicus Journals; Priority Journals

ENTRY MONTH: 200304

ENTRY DATE: Entered STN: 20030331
Last Updated on STN: 20030426
Entered Medline: 20030425

AB We previously determined that all 6 *Staphylococcus aureus* strains with confirmed intermediate-level resistance to glycopeptides (glycopeptide intermediate *S. aureus* [GISA]) from the United States that we tested belonged to **accessory gene regulator** (**agr**) group II. In the present study, we found that 56% of surveyed bloodstream methicillin-resistant *S. aureus* isolates (n = 148) at our hospital were **agr** group II, whereas only 24% of methicillin-susceptible *S. aureus* isolates (n = 33) were **agr** group II (P = .001). Population analysis of genetically engineered **agr**-null and parent wild-type strains of groups I, II, and IV revealed that, when **agr** function is lost, the **agr** group II knockout *S. aureus* was most likely to develop glycopeptide heteroresistance after growth in 1 microg/mL but not 16 microg/mL vancomycin. This strain was unique in showing decreased autolysis after growth in these conditions. This study suggests that some *S. aureus* strains have an intrinsic survival advantage under a glycopeptide selective pressure, which is possibly related to reduced autolysis after exposure to subinhibitory concentrations of glycopeptide.

L76 ANSWER 4 OF 34 MEDLINE on STN
ACCESSION NUMBER: 2003333177 MEDLINE
DOCUMENT NUMBER: PubMed ID: 12865410
TITLE: Salicylic acid attenuates virulence in endovascular infections by targeting global regulatory pathways in *Staphylococcus aureus*.
COMMENT: Comment in: J Clin Invest. 2003 Jul;112(2):149-51. PubMed ID: 12865403
AUTHOR: Kupferwasser Leon Iri; Yeaman Michael R; Nast Cynthia C; Kupferwasser Deborah; Xiong Yan-Qiong; Palma Marco; Cheung Ambrose L; Bayer Arnold S
CORPORATE SOURCE: Division of Infectious Disease, Harbor-UCLA (University of California-Los Angeles), Torrance, California, USA.
CONTRACT NUMBER: AI-37142 (NIAID)
AI-39001 (NIAID)
AI-39108 (NIAID)
SOURCE: Journal of clinical investigation, (2003 Jul) 112 (2) 222-33.
Journal code: 7802877. ISSN: 0021-9738.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Abridged Index Medicus Journals; Priority Journals
ENTRY MONTH: 200308
ENTRY DATE: Entered STN: 20030717
Last Updated on STN: 20030807
Entered Medline: 20030806

AB Aspirin has been previously shown to reduce the in vivo virulence of *Staphylococcus aureus* in experimental endocarditis, through antiplatelet and antimicrobial mechanisms. In the present study, salicylic acid, the major in vivo metabolite of aspirin, mitigated two important virulence phenotypes in both clinical and laboratory *S. aureus* strains: alpha-hemolysin secretion and fibronectin binding in vitro. In addition, salicylic acid reduced the expression of the alpha-hemolysin gene promoter, *hla*, and the fibronectin gene promoter, *fnbA*. Transcriptional analysis, fluorometry, and flow cytometry revealed evidence of salicylic acid-mediated activation of the stress-response gene *sigB*. Expression of the *sigB*-repressible global regulon *sarA* and the global regulon **agr** were also mitigated by salicylic acid, corresponding to the

reduced expression of the hla and fnbA genes in vitro. Studies in experimental endocarditis confirmed the key roles of both sarA and sigB in mediating the antistaphylococcal effects of salicylic acid in vivo. Therefore, aspirin has the potential to be an adjuvant therapeutic agent against endovascular infections that result from *S. aureus*, by downmodulating key staphylococcal global regulons and structural genes in vivo, thus abrogating relevant virulence phenotypes.

L76 ANSWER 5 OF 34 MEDLINE on STN
ACCESSION NUMBER: 2001248100 MEDLINE
DOCUMENT NUMBER: 21189213 PubMed ID: 11292717
TITLE: Subinhibitory clindamycin differentially inhibits transcription of exoprotein genes in *Staphylococcus aureus*.
AUTHOR: Herbert S; Barry P; Novick R P
CORPORATE SOURCE: Program in Molecular Pathogenesis, Skirball Institute, New York University School of Medicine, New York, New York 10016, USA.
CONTRACT NUMBER: RO1-AI30138 (NIAID)
SOURCE: INFECTION AND IMMUNITY, (2001 May) 69 (5) 2996-3003.
Journal code: 0246127. ISSN: 0019-9567.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200105
ENTRY DATE: Entered STN: 20010517
Last Updated on STN: 20010517
Entered Medline: 20010510

AB It has long been known that certain antibiotics, at subinhibitory concentrations, differentially inhibit the synthesis of alpha-hemolysin and other staphylococcal virulence factors. In this report, we show that subinhibitory clindamycin (SBCL) eliminates production of nearly all exoproteins by *Staphylococcus aureus* but has virtually no effect on cytoplasmic proteins. The effect was abolished by a gene conferring resistance to macrolides-lincosamides-streptogramin B, showing that differential inhibition of protein synthesis is responsible; remarkably, however, subinhibitory clindamycin blocked production of several of the individual exoprotein genes, including spa (encoding protein A), hla (encoding alpha-hemolysin), and spr (encoding serine protease), at the level of transcription, suggesting that the primary effect must be differential inhibition of the synthesis of one or more regulatory proteins. In contrast to earlier reports, however, we found that subinhibitory clindamycin stimulates synthesis of coagulase and fibronectin binding protein B, also at the level of transcription. **agr** and **sar** expression was minimally affected by subinhibitory clindamycin. These effects varied from strain to strain and do not seem to be responsible for the effects of subinhibitory clindamycin on the overall exoprotein pattern.

L76 ANSWER 6 OF 34 MEDLINE on STN
ACCESSION NUMBER: 2001285436 MEDLINE
DOCUMENT NUMBER: 21117027 PubMed ID: 11179383
TITLE: Pheromone cross-inhibition between *Staphylococcus aureus* and *Staphylococcus epidermidis*.
AUTHOR: Otto M; Echner H; Voelter W; Gotz F
CORPORATE SOURCE: Mikrobielle Genetik, Abteilung fur Physikalische Biochemie, Universitat Tubingen, Tubingen, Germany..
michael.otto@uni-tuebingen.de
SOURCE: INFECTION AND IMMUNITY, (2001 Mar) 69 (3) 1957-60.
Journal code: 0246127. ISSN: 0019-9567.
PUB. COUNTRY: United States

DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200105
ENTRY DATE: Entered STN: 20010529
Last Updated on STN: 20010529
Entered Medline: 20010524

AB Cross-inhibition by quorum-sensing pheromones between *Staphylococcus aureus* and *Staphylococcus epidermidis* was investigated using all known *S. aureus* **agr** pheromone subgroups. All *S. aureus* subgroups were sensitive towards the *S. epidermidis* pheromone, with the exception of the recently identified subgroup 4. The subgroup 4 pheromone was also the only *S. aureus* pheromone able to inhibit the *S. epidermidis* **agr** response. The close relation of subgroup 4 to subgroup 1 suggests that subgroup 4 might have evolved from subgroup 1 by mutation under the selective pressure of competition with *S. epidermidis*. The competition between *S. aureus* and *S. epidermidis* by means of quorum-sensing cross talk seems to be generally in favor of *S. epidermidis*, which might explain the predominance of *S. epidermidis* on the skin and in infections on indwelling medical devices.

L76 ANSWER 7 OF 34 MEDLINE on STN
ACCESSION NUMBER: 2001541407 MEDLINE
DOCUMENT NUMBER: 21472528 PubMed ID: 11587788
TITLE: *Staphylococcus aureus* and *Staphylococcus epidermidis* peptide pheromones produced by the **accessory gene regulator agr** system.
AUTHOR: Otto M
CORPORATE SOURCE: Rocky Mountain Laboratory, Laboratory of Human Bacterial Pathogenesis, NIAID, NIH, 903 South 4th Street, Hamilton, MT 59840, USA.. motto@niaid.nih.gov
SOURCE: PEPTIDES, (2001 Oct) 22 (10) 1603-8. Ref: 39
Journal code: 8008690. ISSN: 0196-9781.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200112
ENTRY DATE: Entered STN: 20011008
Last Updated on STN: 20020122
Entered Medline: 20011214

AB The **accessory gene regulator (agr)** system of staphylococci regulates the expression of virulence factors in response to cell density. The extracellular signaling molecule encoded by this system is a thiolactone-containing pheromone peptide whose primary sequence varies among staphylococcal strains. A post-translational modification of the peptide is believed to be carried out by an enzyme with a novel function, AgrB. Staphylococcal pheromones show cross-inhibiting properties: Pheromones of self and pheromones of non-self induce and suppress the **agr** response, respectively, and have therefore been proposed as novel anti-staphylococcal drugs. As inhibition of **agr** leads to diminished expression of toxins, but to increased expression of colonization factors and biofilm formation, their therapeutic potential remains yet to be evaluated in depth.

L76 ANSWER 8 OF 34 MEDLINE on STN
ACCESSION NUMBER: 2001284536 MEDLINE
DOCUMENT NUMBER: 20542125 PubMed ID: 11087872
TITLE: Rational design of a global inhibitor of the virulence

response in *Staphylococcus aureus*, based in part on localization of the site of inhibition to the receptor-histidine kinase, AgrC.

AUTHOR: Lyon G J; Mayville P; Muir T W; Novick R P
CORPORATE SOURCE: Laboratory of Synthetic Protein Chemistry, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA.
CONTRACT NUMBER: AI 42783 (NIAID)
GM07739 (NIGMS)
SOURCE: PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, (2000 Nov 21) 97 (24) 13330-5.
Journal code: 7505876. ISSN: 0027-8424.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200105
ENTRY DATE: Entered STN: 20010529
Last Updated on STN: 20010529
Entered Medline: 20010524

AB Two-component signaling systems involving receptor-histidine kinases are ubiquitous in bacteria and have been found in yeast and plants. These systems provide the major means by which bacteria communicate with each other and the outside world. Remarkably, very little is known concerning the extracellular ligands that presumably bind to receptor-histidine kinases to initiate signaling. The two-component **agr** signaling circuit in *Staphylococcus aureus* is one system where the ligands are known in chemical detail, thus opening the door for detailed structure-activity relationship studies. These ligands are short (8- to 9-aa) **peptides** containing a thiolactone structure, in which the alpha-carboxyl group of the C-terminal amino acid is linked to the sulfhydryl group of a cysteine, which is always the fifth amino acid from the C terminus of the **peptide**. One unique aspect of the **agr** system is that **peptides** that activate virulence expression in one group of *S. aureus* strains also inhibit virulence expression in other groups of *S. aureus* strains. Herein, it is demonstrated by switching the receptor-histidine kinase, AgrC, between strains of different **agr** specificity types, that intragroup activation and intergroup inhibition are both mediated by the same group-specific receptors. These results have facilitated the development of a global inhibitor of virulence in *S. aureus*, which consists of a truncated version of one of the naturally occurring thiolactone **peptides**.

L76 ANSWER 9 OF 34 MEDLINE on STN
ACCESSION NUMBER: 2001010995 MEDLINE
DOCUMENT NUMBER: 20461218 PubMed ID: 11004170
TITLE: Population studies of methicillin-resistant and -sensitive *Staphylococcus aureus* strains reveal a lack of variability in the *agrD* gene, encoding a staphylococcal autoinducer **peptide**.
AUTHOR: van Leeuwen W; van Nieuwenhuizen W; Gijzen C; Verbrugh H; van Belkum A
CORPORATE SOURCE: Department of Medical Microbiology and Infectious Diseases, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands.. vanleeuwen@bacl.azr.nl
SOURCE: JOURNAL OF BACTERIOLOGY, (2000 Oct) 182 (20) 5721-9.
Journal code: 2985120R. ISSN: 0021-9193.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals

ENTRY MONTH: 200010
 ENTRY DATE: Entered STN: 20010322
 Last Updated on STN: 20010322
 Entered Medline: 20001026

AB The virulence of *Staphylococcus aureus* is controlled by the **accessory gene regulator (agr)** system, including an extracellular inducer encoded by *agrD*. Variable **agr** PCR restriction fragment length polymorphism (RFLP) patterns of unique *S. aureus* strains (n = 192) were determined for a region comprising *agrD* and parts of the neighboring *agrC* and *agrB* genes. Twelve unique RFLP patterns were identified among *S. aureus* strains in general; these patterns were further specified by sequencing. All sequences could be catalogued in the three current **agr** groups. A major proportion of the *S. aureus* strains belong to **agr** group 1, whereas only 6% of the methicillin-susceptible *S. aureus* strains and 5% of the methicillin-resistant *S. aureus* strains belong to **agr** groups 2 and 3, respectively. The homology between groups varied from 75 to 80%, and within groups it varied from 96 to 100%. Different levels of sequence variability were observed in the different **agr** genes. **agr**-related bacterial interference among colonizing *S. aureus* strains in the noses of persistent and intermittent human carriers was studied. *S. aureus* strains belonging to different **agr** groups were encountered in the same individual. This may suggest that the activity of the *agrD* gene product does not define colonization dynamics, which is further substantiated by the rarity of **agr** group 2 and 3 strains.

L76 ANSWER 10 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 1
 ACCESSION NUMBER: 2002:688473 CAPLUS
 DOCUMENT NUMBER: 137:222001
 TITLE: Blocking expression of virulence factors in *Staphylococcus aureus* with peptide compositions
 INVENTOR(S): Novick, Richard P.; Ji, Guangyong; Beavis, Ronald
 PATENT ASSIGNEE(S): New York University, USA
 SOURCE: U.S., 12 pp., Cont.-in-part of U.S. Ser. No. 651,226, abandoned.
 CODEN: USXXAM
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 3
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6447786	B1	20020910	US 1997-861476	19970522
US 2003078378	A1	20030424	US 2002-201444	20020723
PRIORITY APPLN. INFO.:			US 1994-318499	B2 19941004
			US 1996-651226	B2 19960522
			US 1997-861476	A1 19970522

AB This invention provides peptides which inhibit **agr** transcription in *S. aureus* and thereby block the expression of virulence factors in *S. aureus*, pharmaceutical compns. comprising these peptides, as well as methods for treating or preventing an infection or disease caused by *S. aureus* using the peptides of the present invention.

REFERENCE COUNT: 15 THERE ARE 15 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L76 ANSWER 11 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN
 ACCESSION NUMBER: 2004:41215 CAPLUS
 TITLE: Preparation of nocathiacin O-derivatives
 INVENTOR(S): Hudyma, Thomas W.; Zheng, Xiaofan; Naidu, B.

Narasimhulu; Sorenson, Margaret E.; Regueiro-Ren,
Alicia; Connolly, Timothy P.; Matiskella, John D.;
Kim, Oak K.; Zhang, Yunhui; Springer, Dane M.;
Goodrich, Jason; Ueda, Yasutsugu; Bronson, Joanne J.
PATENT ASSIGNEE(S): Bristol-Myers Squibb Company, USA
SOURCE: PCT Int. Appl., 53 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2004004646	A2	20040115	WO 2003-US21012	20030702
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ			
RW:	GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG			

US 2004018963 A1 20040129 US 2002-189710 20020703

PRIORITY APPLN. INFO.: US 2002-189710 A 20020703

AB The invention provides compds. I [Z = NH₂ or NHC(:CH₂)CONH₂; R₁ = 1-methyl-4-imidazolylsulfonyl, tripyrrolinophosphonium, H, phosphonyl, acyl, or carbamoyl groups, (CH₂CH₂O)1-6Me, (un)substituted alkyl, alkenyl, alkynyl; R₂ = hydroxy-functionalized tetrahydropyran derivs., H, phosphonyl, acyl, or carbamoyl groups, SO₃H, (CH₂CH₂O)1-6H, (CH₂CH₂O)1-6Me, (un)substituted alkyl, alkenyl, alkynyl] having potent antibiotic activity, including activity towards Gram-pos. bacteria and mycobacteria. Thus, I (Z = NHC(:CH₂)CONH₂, R₁ = Me, R₂ = H) was prepared by O-methylation of nocathiacin I and showed MIC = 0.015, 0.003, and 0.015 µg/mL assayed against Staphylococcus aureus, Streptococcus pneumoniae, and Enterococcus faecalis, resp.

L76 ANSWER 12 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2003:1007003 CAPLUS

DOCUMENT NUMBER: 140:59934

TITLE: Synthesis of cyclic peptide macrolides for use in the treatment and prevention of bacterial infection

INVENTOR(S): Lampe, Thomas; Adelt, Isabelle; Beyer, Dieter; Brunner, Nina; Endermann, Rainer; Ehlert, Kerstin; Kroll, Hein-Peter; Von Nussbaum, Franz; Raddatz, Siegfried; Rudolph, Joachim; Schiffer, Guido; Schumacher, Andreas; Cancho-Grande, Yolanda; Michels, Martin; Weigand, Stefan

PATENT ASSIGNEE(S): Bayer Healthcare A.-G., Germany

SOURCE: PCT Int. Appl., 190 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: German

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
------------	------	------	-----------------	------

WO 2003106480 A1 20031224 WO 2003-EP6078 20030610

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG

DE 10226921 A1 20031224 DE 2002-10226921 20020617

PRIORITY APPLN. INFO.:

DE 2002-10226921 A 20020617

AB The invention relates to antibacterial amide macrocycles, e.g. (I), to methods for the production thereof, and to the use of the same for producing pharmaceuticals for the treatment and/or prophylaxis of illness, especially bacterial infections. Title compds. were synthesized beginning with salicylaldehyde, which was 5-iodinated, O-protected, reduced to the hydroxymethyl, brominated on the CH₂ group, and coupled with di-Et 2-tert-butoxycarbonylaminomalonate, which, after decarboxylation and deesterification, gave the (DL)-N-Boc-protected 2'-benzyloxy-5'-iodo-phenylalanine (II). II was resolved into its pure D- and L-enantiomers; the L-II was protected as the N-Cbz derivative, then esterified with 2-(trimethylsilyl)ethanol, then reacted with (III) (prepared from II and 4,4,4',4',5,5,5',5'-octamethyl-2,2'-bi-1,3,2-dioxaborolane) to give biphenyl compound (IV). In a sep. reaction, (V) was prepared from the corresponding L-ornithine tert-Bu ester, the lactone opened and the alc. protected as the tert-butyldimethylsilyl derivative, and reacted with biphenyl IV, to give, after deprotection and amide formation, I as the dihydrochloride salt. In in vitro tests, using *S. aureus*, *E. faecalis*, *B. catarrhalis*, and *E. coli* strains, I had min. blood concentration effective ranges of 0.2-6.25 µM.

REFERENCE COUNT: 5 THERE ARE 5 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L76 ANSWER 13 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2003:521316 CAPLUS

DOCUMENT NUMBER: 139:277145

TITLE: Side-chain-to-tail thiolactone peptide inhibitors of the staphylococcal quorum-sensing system

AUTHOR(S): Scott, R. John; Lian, Lu-Yun; Muharram, S. Hanna; Cockayne, Alan; Wood, Stewart J.; Bycroft, Barrie W.; Williams, Paul; Chan, Weng C.

CORPORATE SOURCE: School of Pharmaceutical Sciences, University of Nottingham, Nottingham, NG7 2RD, UK

SOURCE: Bioorganic & Medicinal Chemistry Letters (2003), 13(15), 2449-2453

CODEN: BMCLE8; ISSN: 0960-894X

PUBLISHER: Elsevier Science B.V.

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The expression of many staphylococcal virulence factors are regulated by the **agr** locus via a two-component signal transduction system (TCSTS), which is activated in response to a secreted autoinducer peptide (AIP). By exploiting the unique chemical architecture of the naturally occurring AIP-1, several potent inhibitors of staphylococcal TCSTS were designed and synthesized using either a linear or branched solid-phase approach. These prepared inhibitors I [R₁ = CH(NH₂)(S)-CH₂-C₆H₄OH, R₂ = Me, or CH₂Me; R₁ = (CH₂)₃OC₆H₄CH₂Ph, or (CH₂)₃OC₆H₄COPh, R₂ = Me] are competitive binders and contain the crucial 16-membered side-chain-to-tail

thiolactone peptide pharmacophore.

REFERENCE COUNT: 30 THERE ARE 30 CITED REFERENCES AVAILABLE FOR THIS
RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L76 ANSWER 14 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2003:171160 CAPLUS

DOCUMENT NUMBER: 138:332795

TITLE: Transcriptional induction of the penicillin-binding
protein 2 gene in *Staphylococcus aureus* by cell
wall-active antibiotics oxacillin and vancomycin
AUTHOR(S): Boyle-Vavra, Susan; Yin, Shaohui; Challapalli,
Mamatha; Daum, Robert S.

CORPORATE SOURCE: Department of Pediatrics, The University of Chicago,
Chicago, IL, 60637, USA

SOURCE: Antimicrobial Agents and Chemotherapy (2003), 47(3),
1028-1036

CODEN: AMACCQ; ISSN: 0066-4804

PUBLISHER: American Society for Microbiology

DOCUMENT TYPE: Journal

LANGUAGE: English

AB We found an increased abundance of pbpB-specific transcripts in vancomycin intermediate-resistant *Staphylococcus aureus* (VISA) isolates compared with that found in paired, genetically identical, susceptible isolates. This difference in expression cannot be explained by differences in the pbpB promoter sequence. Since the factors controlling pbpB gene expression have remained largely unexplored, various conditions that might affect pbpB transcript abundance were examined. In both vancomycin-susceptible and VISA strains, pbpB expression varied with the growth phase, with the highest abundance of pbpB-specific transcripts detected during mid-log phase. Interestingly, both vancomycin and oxacillin were able to induce pbpB transcription above a constitutive level. When vancomycin was absent, one of the three pbpB-specific transcripts that were usually faintly detected in non-VISA strains was more readily detected in VISA strains during mid-log but not stationary phase. This transcript was enhanced in non-VISA strains by vancomycin induction. Gel shift assays indicated that an increased amount of the putative transcription factor that binds to both P1 and P1' promoter regions is present in the cytosol of vancomycin-induced cells. Neither the SigB sigma factor nor the quorum-sensing *agr* locus was required for growth phase-variable pbpB expression or transcriptional induction of pbpB by vancomycin or oxacillin. Also, MecI, MecR1, BlaI, and BlaR1, regulatory proteins that mediate β -lactam-inducible expression of *mecA* and β -lactamase, were not required for antibiotic induction of pbpB transcription. These data support the idea that pbpB expression is modulated by a trans-acting factor in response to the presence of the cell wall-active antibiotics vancomycin and oxacillin.

REFERENCE COUNT: 51 THERE ARE 51 CITED REFERENCES AVAILABLE FOR THIS
RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L76 ANSWER 15 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2002:869046 CAPLUS

DOCUMENT NUMBER: 137:363038

TITLE: Antimicrobial cyclic peptides, compositions containing
them, and screening methods

INVENTOR(S): Ghadiri, M. Reza; Kim, Hui-Sun; Fernandez-Lopez, Sara;
Wilcoxon, Keith

PATENT ASSIGNEE(S): The Scripps Research Institute, USA

SOURCE: PCT Int. Appl., 240 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002090503	A2	20021114	WO 2002-US14329	20020506
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG				
WO 2003093300	A2	20031113	WO 2003-US14240	20030506
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG				

PRIORITY APPLN. INFO.: US 2001-288990P P 20010504
WO 2002-US14329 A 20020506

AB The invention provides antimicrobial agents and compns. that include cyclic peptides having an amino sequence of alternating D-and L- α -amino acids. Alternatively, the cyclic peptides are made from β -amino acids. Methods for identifying and evaluating antimicrobial cyclic peptides are also provided.

L76 ANSWER 16 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2002:676289 CAPLUS

DOCUMENT NUMBER: 137:211942

TITLE: Drug design against drug resistant mutants using **directed** evolution and target protein conformation changes

INVENTOR(S): Stevens, Raymond C.; Orenica, Maria C.; Yoon, Jun S.; Hanson, Michael A.

PATENT ASSIGNEE(S): The Scripps Research Institute, USA

SOURCE: PCT Int. Appl., 82 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002068933	A2	20020906	WO 2002-US6238	20020227
WO 2002068933	A3	20021121		
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ,				

UA, UG, US, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU,
TJ, TM
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH,
CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR,
BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG

PRIORITY APPLN. INFO.: US 2001-272248P P 20010228

AB The present invention provides methods for identifying new drugs and potential inhibitors and modulators of drug-resistant variants of a target protein of a drug of interest. A drug-resistant variant according to the invention has at least one mutation resulting in a structural change, an activity change or a stability change as compared to the target protein. Such variants would include natural variants such as those encountered in the clinic, but preferably variants are selected by directed evolution methodol. The present invention relates to methods for designing new drugs useful against drug-resistant bacterial cells, viruses, mammalian cells and the like. The method involves identifying a target protein of the drug, selecting for drug-resistant variants that have an altered target protein (variant protein) by directed evolution, determining the three dimensional structure of the target and variant proteins and designing a new drug that can be effective against at least one drug-resistant variant. The present invention can be used to predict future mutations that lead to drug resistance and the type of drugs that are effective to combat such resistance.

L76 ANSWER 17 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2002:71893 CAPLUS

DOCUMENT NUMBER: 136:129034

TITLE: Antimicrobial sulfonamide derivatives of lipopeptide antibiotics

INVENTOR(S): Curran, William V.; Leese, Richard A.; Jarolmen, Howard; Borders, Donald B.

PATENT ASSIGNEE(S): Intrabiotics Pharmaceuticals, Inc., USA

SOURCE: PCT Int. Appl., 67 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 6

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002005837	A1	20020124	WO 2001-US22352	20010717
W:				
AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW:				
GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG				
EP 1311281	A1	20030521	EP 2001-957162	20010717
R:				
AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL, TR				

PRIORITY APPLN. INFO.: US 2000-219059P P 20000717

US 2000-220950P P 20000726

US 2001-760328 A 20010112

WO 2001-US22352 W 20010717

OTHER SOURCE(S): MARPAT 136:129034

AB The invention provides antimicrobial sulfonamide derivs. of lipopeptide antibiotics, pharmaceutical compns. of antimicrobial sulfonamide derivs.,

methods for making antimicrobial sulfonamide derivs., methods for inhibiting microbial growth with antimicrobial sulfonamide derivs., and methods for treating or preventing microbial infections in a subject with antimicrobial sulfonamide derivs. Antimicrobial sulfonamide derivs. are generally an amino core antibiotic that has been further modified with a lipophilic sulfonyl group.

REFERENCE COUNT: 6 THERE ARE 6 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L76 ANSWER 18 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2002:907161 CAPLUS

DOCUMENT NUMBER: 138:13500

TITLE: Superantigen-glycolipid conjugates loaded onto antigen presenting cells for adoptive immunotherapy of neoplastic and infectious diseases

INVENTOR(S): Terman, David S.

PATENT ASSIGNEE(S): USA

SOURCE: U.S. Pat. Appl. Publ., 167 pp.

CODEN: USXXCO

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2002177551	A1	20021128	US 2001-870759	20010530

PRIORITY APPLN. INFO.: US 2000-208128P P 20000531

AB The present invention comprises compns. and methods for treating a tumor or neoplastic disease in a host. The methods employ conjugates comprising superantigen polypeptides, nucleic acids with other structures that preferentially bind to tumor cells and are capable of inducing apoptosis. Also provided are superantigen-glycolipid conjugates and vesicles that are loaded onto antigen presenting cells to activate both T cells and NKT cells. Cell-based vaccines comprise tumor cells engineered to express a superantigen along with glycolipids products which, when expressed, render the cells capable of eliciting an effective anti-tumor immune response in a mammal into which these cells are introduced. Included among these compns. are tumor cells, hybrid cells of tumor cells and accessory cells, preferably dendritic cells. Also provided are tumoricidal T cells and NKT cells devoid of inhibitory receptors or inhibitory signaling motifs which are hyperresponsive to the the above compns. and lipid-based tumor associated antigens that can be administered for adoptive immunotherapy of cancer and infectious diseases.

L76 ANSWER 19 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2002:874013 CAPLUS

DOCUMENT NUMBER: 138:106988

TITLE: Synthesis of Novel Nocathiacin-Class Antibiotics. Condensation of Glycolaldehyde with Primary Amides and Tandem Reductive Amination of Amadori-Rearranged 2-Oxoethyl Intermediates.

AUTHOR(S): Hrciar, Peter; Ueda, Yasutsugu; Huang, Stella; Leet, John E.; Bronson, Joanne J.

CORPORATE SOURCE: Pharmaceutical Research Institute, Bristol-Myers Squibb Co., Wallingford, CT, 06492, USA

SOURCE: Journal of Organic Chemistry (2002), 67(25), 8789-8793
CODEN: JOCEAH; ISSN: 0022-3263

PUBLISHER: American Chemical Society

DOCUMENT TYPE: Journal

LANGUAGE: English

OTHER SOURCE(S): CASREACT 138:106988

AB Nocathiacin I, I [R = NHC(:CH₂)CONH₂], and nocathiacin IV, I (R = NH₂), are novel indole-containing thiazolyl cyclic peptide antibiotics, which exhibit potent activity against key Gram-pos. bacterial pathogens, including multi drug-resistant *Staphylococcus aureus*, *Streptococcus pneumoniae*, and *Enterococcus faecium*. New nocathiacins I (R = NHCH₂CH₂NR₁R₂; NR₁R₂ = morpholine; R₁ = R₂ = CH₂CH₂OH; R₁ = Et, R₂ = CH₂CH₂OH; R₁ = Me, R₂ = CH₂CH₂OH; R₁ = R₂ = Et; R₁ = Et, R₂ = CH₂CO₂H) were prepared from nocathiacin IV by condensation with glycolaldehyde dimer followed by tandem reductive amination of the intermediate with NHR₁R₂. This transformation readily tolerated the complex architecture of nocathiacin IV and allowed selective incorporation of water-solubilizing groups to the primary amide in nocathiacin IV without needing the presence of protecting groups.

REFERENCE COUNT: 25 THERE ARE 25 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L76 ANSWER 20 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2002:853427 CAPLUS

DOCUMENT NUMBER: 138:304509

TITLE: The synthesis of a novel binaphthyl-based cyclic peptoid with anti-bacterial activity

AUTHOR(S): Bremner, John B.; Coates, Jonathan A.; Coghlan, Daniel R.; David, Dorothy M.; Keller, Paul A.; Pyne, Stephen G.

CORPORATE SOURCE: Department of Chemistry, University of Wollongong, Wollongong, 2522, Australia

SOURCE: New Journal of Chemistry (2002), 26(11), 1549-1551
CODEN: NJCHE5; ISSN: 1144-0546

PUBLISHER: Royal Society of Chemistry

DOCUMENT TYPE: Journal

LANGUAGE: English

OTHER SOURCE(S): CASREACT 138:304509

AB The novel cyclic peptoid I, based upon a 1,1'-binaphthyl scaffold and a bridging tripeptide moiety, was synthesized utilizing a ring-closing metathesis reaction. The individual major and minor diastereomers of compound I were shown to have promising anti-bacterial activity against *Staphylococcus aureus* with inhibition of microbial growth at 17 µL-1m (MIC) and 31 µg ml⁻¹, resp.

REFERENCE COUNT: 25 THERE ARE 25 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L76 ANSWER 21 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2002:287801 CAPLUS

DOCUMENT NUMBER: 137:185805

TITLE: Synthesis and evaluation of vancomycin and vancomycin aglycon analogues that bear modifications in the residue 3 asparagine

AUTHOR(S): McAtee, J. Jeffrey; Castle, Steven L.; Jin, Qing; Boger, Dale L.

CORPORATE SOURCE: Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA, 92037, USA

SOURCE: Bioorganic & Medicinal Chemistry Letters (2002), 12(9), 1319-1322
CODEN: BMCLE8; ISSN: 0960-894X

PUBLISHER: Elsevier Science Ltd.

DOCUMENT TYPE: Journal

LANGUAGE: English

OTHER SOURCE(S): CASREACT 137:185805

AB The synthesis and biol. evaluation of a set of residue 3 analogs of

vancomycin and its aglycon are described. These investigations follow from the promising biol. activity of a protected and synthetically modified vancomycin aglycon analog in which the asparagine side chain was modified to possess a nitrile, rather than a carboxamide. Although this modification typically was detrimental to antimicrobial activity, hydrophobic vancomycin aglycon analogs that lack a lipid anchor as well as the disaccharide are detailed that exhibit unusual potency against VanB, but not VanA, resistant bacteria.

REFERENCE COUNT: 22 THERE ARE 22 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L76 ANSWER 22 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2001:152655 CAPLUS

DOCUMENT NUMBER: 134:193744

TITLE: Preparation of macrocyclic compounds for treating bacterial infections

INVENTOR(S): Jefferson, Elizabeth; Swayze, Eric Edward

PATENT ASSIGNEE(S): Isis Pharmaceuticals, Inc., USA

SOURCE: PCT Int. Appl., 98 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001014346	A1	20010301	WO 2000-US22871	20000818
W:		AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM		
RW:		GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG		
US 6660832	B1	20031209	US 1999-378529	19990820

PRIORITY APPLN. INFO.:

US 1999-378529 A 19990820

OTHER SOURCE(S): MARPAT 134:193744

AB Macrocyclic compds. I [X = O, NH, S; R1-R4 = H, amino or an amino acid side chain; R5 = H, OH, CO₂H, halo, SH, cyano, amino, an electron withdrawing group, alkoxy, C(O)NH₂, -C(O)NHR₆, C(O)-(amino acid)1-4, -C(O)OR₆, CH₂OH, -CH₂OR₆, NHC(O)R₇, or -NH-(amino acid)1-4; R₆, R₇ = alkyl (optionally substituted with OH, halogen, CO₂H, cyano, amino, amidine, guanidine, urea, or a nucleobase) or aryl, aralkyl, heterocyclyl, or heterocyclalkyl optionally substituted with OH, halogen, CO₂H, oxo, cyano, amino, amidine, guanidine, or urea (the substituent may also be C1-4 alkyl in the case of R₇)] are useful for therapeutic and prophylactic treatment of bacterial infection in mammals. Solid phase synthetic procedures are given for the synthesis of a library of macrocyclic compds. using KanTM reactors. Claimed compds. include macrocycle II. Min. inhibitory concentration and coupled bacterial transcription/translation data are tabulated.

L76 ANSWER 23 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2001:225071 CAPLUS

DOCUMENT NUMBER: 135:16454

TITLE: Pseudodestruxins A and B: New Cyclic Depsipeptides from the Coprophilous Fungus Nigrosabulum globosum

AUTHOR(S): Che, Yongsheng; Swenson, Dale C.; Gloer, James B.;

CORPORATE SOURCE: Koster, Brenda; Malloch, David
 Department of Chemistry, University of Iowa, Iowa
 City, IA, 52242, USA

SOURCE: Journal of Natural Products (2001), 64(5), 555-558
 CODEN: JNPRDF; ISSN: 0163-3864

PUBLISHER: American Chemical Society

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Pseudodestruxins A (I) and B (II), two new cyclic peptides, have been isolated from cultures of the coprophilous fungus *Nigrosabulum globosum*. The structure of pseudodestruxin A was elucidated using 2D NMR techniques and confirmed by single-crystal X-ray diffraction anal. The structure of pseudodestruxin B was assigned by comparing its NMR and FABMS data with those of pseudodestruxin A. The known compds. ascochlorin and 5-chlorocollectorin B were also isolated from *N. globosum*. Although Pseudodestruxins A and B display antibacterial effects, ascochlorin was found to be responsible for the antifungal activity of the crude extract

REFERENCE COUNT: 12 THERE ARE 12 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L76 ANSWER 24 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 1999:297439 CAPLUS

DOCUMENT NUMBER: 130:297010

TITLE: Preparation of cyclic peptides having broad spectrum antimicrobial activity

INVENTOR(S): Chang, Conway; Gu, Leo; Chen, Jie

PATENT ASSIGNEE(S): Intrabiotics Pharmaceuticals, Inc., USA

SOURCE: PCT Int. Appl., 167 pp.
 CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9921879	A1	19990506	WO 1997-US19557	19971027
W: AL, AM, AU, AZ, BA, BB, BG, BR, BY, CA, CN, CU, CZ, EE, GE, GH, HU, ID, IL, IS, JP, KG, KP, KR, KZ, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UZ, VN, YU, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM RW: GH, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG				
AU 9851535	A1	19990517	AU 1998-51535	19971027
PRIORITY APPLN. INFO.: WO 1997-US19557 19971027				
AB The present invention provides cyclic peptides I [m = 0-2, n = 0-1, with the proviso that when m = 2, n = 0; X21, X22, X24, X25, X27, X28 are each independently present or absent; X7 and X4 are either both present or both absent; X8 and X3 are either both present or both absent; X2-X5, X7, X8, X13, X14, X16-X19, X21, X22, X27, X28 independently = hydrophobic amino acid, hydrophilic amino acid, small amino acid, with provisos (i) when X2 = hydrophobic amino acid, X7, X14, X19, X21, and X28 independently = hydrophobic amino acid or small amino acid and X3, X8, X13, X18, X22 and X27 independently = hydrophilic amino acid or small amino acid and (ii) when X2 = hydrophilic amino acid X7, X14, X19, X21, and X28 independently = hydrophilic amino acid or small amino acid and X3, X8, X13, X18, X22 and X27 independently = hydrophobic amino acid or small amino acid; X23-X26 taken together = loop; Z1, Z6, Z5, Z20 independently = hydrophilic amino acid, small amino acid, cysteine-like amino acid; X9-X12 taken together = β -turn; at least one of X9-X12, X23-X26 = basic amino acid; and				

wherein the peptide has net pos. charge at physiol. pH] comprising and amphiphilic antiparallel β -sheet region, a loop region, and a β -turn region having broad spectrum antimicrobial activity. The peptides exhibit improved efficacy, bioavailability and/or serum half-life as compared with non-cyclized analogs. Thus, cystine-containing cyclopeptide II inhibited *Pseudomonas aeruginosa* with MIC = 8 μ g/mL and methicillin-resistant *Staphylococcus aureus* with MIC = 2 μ g/mL compared to 32 μ g/mL against both bacteria for the uncyclized peptide. In addition, II showed increased activities after 15 min and 120 min relative to the uncyclized peptide.

REFERENCE COUNT: 1 THERE ARE 1 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L76 ANSWER 25 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 1998:719131 CAPLUS

DOCUMENT NUMBER: 130:4085

TITLE: Preparation of peptides with bactericidal activity and endotoxin neutralizing activity for gram negative bacteria

INVENTOR(S): Gray, Beulah H.; Haseman, Judith R.; Mayo, Kevin H.

PATENT ASSIGNEE(S): Regents of the University of Minnesota, USA

SOURCE: U.S., 56 pp., Cont.-in-part of U.S. Ser. No. 218,026.

CODEN: USXXAM

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 4

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 5830860	A	19981103	US 1996-653632	19960524
US 5786324	A	19980728	US 1994-218026	19940324
WO 9744354	A2	19971127	WO 1997-US8944	19970523
W: CA, JP, US, US				
RW: AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
EP 939766	A2	19990908	EP 1997-928665	19970523
R: DE, FR, GB, IT, SE				
JP 2000511892	T2	20000912	JP 1997-542843	19970523
US 6486125	B1	20021126	US 1999-194296	19991015
US 2003153502	A1	20030814	US 2002-300083	20021120

PRIORITY APPLN. INFO.:
 US 1994-218026 A2 19940324
 US 1996-653632 A2 19960524
 US 1996-671487 A2 19960627
 WO 1997-US8944 W 19970523
 US 1999-194296 A3 19991015

AB The invention provides biol. active peptides derived from or corresponding to regions of a bactericidal permeability increasing factor (B/PI). The peptides are preferably about 10 to 100 amino acids long and have bactericidal and/or endotoxin neutralizing activity. The peptides can be prepared by automated protein synthesis or by recombinant DNA methods. The peptides are useful in methods to treat and prevent bacterial infection in the body and on surfaces. The peptides are also useful to treat endotoxin shock and have endotoxin neutralizing activity.

REFERENCE COUNT: 35 THERE ARE 35 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L76 ANSWER 26 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 1998:508906 CAPLUS

DOCUMENT NUMBER: 129:156917

TITLE: Synthetic peptides with bactericidal activity and endotoxin neutralizing activity for gram negative

INVENTOR(S): bacteria and methods for their use
 Gray, Beulah; Haseman, Judith R.; Mayo, Kevin
 PATENT ASSIGNEE(S): Regents of the University of Minnesota, USA
 SOURCE: U.S., 46 pp.
 CODEN: USXXAM
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 4
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 5786324	A	19980728	US 1994-218026	19940324
US 5830860	A	19981103	US 1996-653632	19960524

PRIORITY APPLN. INFO.: US 1994-218026 A2 19940324

AB The invention provides biol. active peptides derived from or corresponding to regions of a bactericidal permeability increasing factor (B/PI). The peptides are about 10 to 100 amino acids long and have bactericidal and/or endotoxin neutralizing activity. The peptides can be prepared by automated DNA synthesis or by recombinant DNA methods. The peptides are useful in methods to treat and prevent bacterial infection in the body and on surfaces. The peptides are also useful to treat endotoxin shock.

REFERENCE COUNT: 76 THERE ARE 76 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L76 ANSWER 27 OF 34 CAPLUS COPYRIGHT 2004 ACS on STN
 ACCESSION NUMBER: 1997:414155 CAPLUS
 DOCUMENT NUMBER: 127:31334
 TITLE: Kawaguchipectin B, an antibacterial cyclic undecapeptide from the cyanobacterium *Microcystis aeruginosa*
 AUTHOR(S): Ishida, Keishi; Matsuda, Hisashi; Murakami, Masahiro; Yamaguchi, Katsumi
 CORPORATE SOURCE: Laboratory of Marine Biochemistry Graduate School of Agricultural Life Sciences, University of Tokyo, Tokyo, 113, Japan
 SOURCE: Journal of Natural Products (1997), 60(7), 724-726
 CODEN: JNPRDF; ISSN: 0163-3864
 PUBLISHER: American Chemical Society
 DOCUMENT TYPE: Journal
 LANGUAGE: English

AB Kawaguchipectin B, an antibacterial cyclic undecapeptide, was isolated from the cultured cyanobacterium *Microcystis aeruginosa* (NIES-88). Its structure was elucidated as (I) on the basis of 2D NMR data and chemical degradation. I inhibited the growth of the Gram-pos. bacterium *Staphylococcus aureus* at a concentration of 1 µg/mL (MIC).

L76 ANSWER 28 OF 34 BIOSIS COPYRIGHT 2004 BIOLOGICAL ABSTRACTS INC. on STN
 ACCESSION NUMBER: 2004:22398 BIOSIS
 DOCUMENT NUMBER: PREV200400009801
 TITLE: An **engineered** multidomain **bactericidal** peptide as a model for targeted antibiotics against specific bacteria.
 AUTHOR(S): Qiu, Xiao-Qing; Wang, He; Lu, Xiao-Fong; Zhang, Jie; Li, Sheng-Fu; Cheng, Gang; Wan, Lin; Yang, Li; Zuo, Jun-Yong; Zhou, Yu-Qi; Wang, Hai-Yun; Cheng, Xin; Zhang, Su-Hua; Ou, Zheng-Rong; Zhong, Zi-Cheng; Cheng, Jing-Qiu; Li, You-Ping; Wu, George Y. [Reprint Author]
 CORPORATE SOURCE: Division of Gastroenterology-Hepatology, University of Connecticut Health Center, Farmington, CT, 06030-1845, USA
 qiu@mrsa.com.cn; wu@nso.uchc.edu

SOURCE: Nature Biotechnology, (December 2003) Vol. 21, No. 12, pp. 1480-1485. print.
ISSN: 1087-0156 (ISSN print).
DOCUMENT TYPE: Article
LANGUAGE: English
OTHER SOURCE: DDBJ-AF001782; EMBL-AF001782; GenBank-AF001782;
DDBJ-AF001783; EMBL-AF001783; GenBank-AF001783;
DDBJ-U85097; EMBL-U85097; GenBank-U85097
ENTRY DATE: Entered STN: 24 Dec 2003
Last Updated on STN: 24 Dec 2003

AB We constructed a peptide consisting of a staphylococcal AgrD1 pheromone fused to the channel-forming domain of colicin Ia and named it pheromonycin. This fusion peptide had **bactericidal** effects against methicillin-sensitive and methicillin-resistant **Staphylococcus aureus** (MSSA and MRSA, respectively), but not against *Staphylococcus epidermidis* or *Streptococcus pneumoniae*. Growth rates, vital staining and colony forming unit (CFU) counts showed that pheromonycin did not merely suppress growth but killed **S. aureus** cells. The specificity of pheromonycin was shown by the absence of **bactericidal** effects against an **accessory gene regulator (agr)** locus knockout of **S. aureus**, and a dose-dependent inhibition of the **bactericidal** effects of pheromonycin by competition with corresponding free AgrD pheromone. In vivo, all pheromonycin-treated mice survived administration of MRSA that was lethal to controls. No toxicity was detectable in human liver or renal cells in culture, or in livers, kidneys or spleens of pheromonycin-treated mice. The results suggest that these types of chimeric peptides may be of value as antibiotics against specific bacterial infections.

L76 ANSWER 29 OF 34 WPIDS COPYRIGHT 2004 THOMSON DERWENT on STN
ACCESSION NUMBER: 2003-606005 [57] WPIDS
CROSS REFERENCE: 1996-209319 [21]; 1998-018426 [02]; 2003-074097 [07]
DOC. NO. CPI: C2003-165005
TITLE: New purified peptide that inhibits **agr-rnaIII** transcription in **Staphylococcus aureus**, useful for treating or preventing a staphylococcal infection or disease.
DERWENT CLASS: B04
INVENTOR(S): BEAVIS, R; JI, G; NOVICK, R P
PATENT ASSIGNEE(S): (BEAV-I) BEAVIS R; (JIGG-I) JI G; (NOVI-I) NOVICK R P
COUNTRY COUNT: 1
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
US 2003078378	A1	20030424	(200357)*		12

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
US 2003078378	A1 CIP of	US 1994-318499	19941004
	CIP of	US 1996-651226	19960522
	Cont of	US 1997-861476	19970522
		US 2002-201444	20020723

FILING DETAILS:

PATENT NO	KIND	PATENT NO

US 2003078378 A1 Cont of

US 6447786

PRIORITY APPLN. INFO: US 1997-861476 19970522; US 1994-318499
19941004; US 1996-651226 19960522; US
2002-201444 20020723

AB US2003078378 A UPAB: 20030906

NOVELTY - A purified peptide, which inhibits *agr*-rnaIII transcription in *Staphylococcus aureus*, is new. The peptide contains 6-12 amino acids in length and comprises amino acid 28 from the AgrD region of a Staphylococcibacterium, or its analogue.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for:

(1) a peptide composition comprising the purified peptide cited above and a pharmaceutical carrier; and

(2) treating or preventing an infection or disease caused by *S. aureus* in a subject comprising administering to the subject the purified peptide.

ACTIVITY - **Antibacterial.**

No biological data given.

MECHANISM OF ACTION - Virulence Factor Inhibitor.

USE - The purified peptides and compositions are useful for treating or preventing an infection or disease caused by a Staphylococci bacterium. Dwg.0/4

L76 ANSWER 30 OF 34 WPIDS COPYRIGHT 2004 THOMSON DERWENT on STN

ACCESSION NUMBER: 2002-681366 [73] WPIDS

CROSS REFERENCE: 2002-170774 [22]

DOC. NO. CPI: C2002-192257

TITLE: New **cyclic peptides**, useful for treating *Staphylococcus aureus* infections.

DERWENT CLASS: A96 B04

INVENTOR(S): BEAVIS, R; JI, G; MAYVILLE, P; MUIR, T W; NOVICK, R P

PATENT ASSIGNEE(S): (BEAV-I) BEAVIS R; (JIGG-I) JI G; (MAYV-I) MAYVILLE P;
(MUIR-I) MUIR T W; (NOVI-I) NOVICK R P

COUNTRY COUNT: 1

PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
US 2002077453	A1	20020620	(200273)*		18

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
US 2002077453	A1	Provisional	US 1998-90402P 19980624
		Cont of	US 1999-339511 19990624
			US 2001-32950 20011227

PRIORITY APPLN. INFO: US 1998-90402P 19980624; US 1999-339511
19990624; US 2001-32950 20011227

AB US2002077453 A UPAB: 20021113

NOVELTY - A **cyclic peptide** is new.

DETAILED DESCRIPTION - A **cyclic peptide** of formula (I) is new.

X = amino acid, amino acid analog, peptidomimetic or non-amide isostere;

Z' = synthetic or biosynthetic amino acid;

R = O, N or C;

n = 0-10; and

y = 1-10.

INDEPENDENT CLAIMS are included for:

- (1) a **cyclic peptide** of formula $H_2N-X_n-Z'-X_y-COOH$
 (II) with a cyclic bond between Z' and COOH other than a thioester bond;
 (2) compositions comprising the peptides;
 (3) method for treating **S. aureus** infection
 comprises administration of the composition;
 (4) method for preparing (I).

ACTIVITY - **Antibacterial**.

MECHANISM OF ACTION - **Agr** response inhibitor.

USE - The peptides are useful for treating **S. aureus** infections (claimed).
 Dwg.0/3

L76 ANSWER 31 OF 34 WPIDS COPYRIGHT 2004 THOMSON DERWENT on STN
 ACCESSION NUMBER: 2002-170774 [22] WPIDS
 CROSS REFERENCE: 2002-681366 [73]
 DOC. NO. CPI: C2002-052666
 TITLE: Novel synthetic, **cyclic AgrD**-autoinducing
peptide for bacterial interference and for
 treating **Staphylococcus aureus**
 infection in a subject.

DERWENT CLASS: B03 B04 D16
 INVENTOR(S): BEAVIS, R; JI, G; MAYVILLE, P; MUIR, T W; NOVICK, R P
 PATENT ASSIGNEE(S): (UYNY) UNIV NEW YORK STATE; (UYRQ) UNIV ROCKEFELLER
 COUNTRY COUNT: 1
 PATENT INFORMATION:

Sellward →

PATENT NO	KIND	DATE	WEEK	LA	PG
US 6337385	B1	20020108	(200222)*		18

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
US 6337385	B1 Provisional	US 1998-90402P	19980624
		US 1999-339511	19990624

PRIORITY APPLN. INFO: US 1998-90402P 19980624; US 1999-339511
 19990624

AB US 6337385 B UPAB: 20021118

NOVELTY - A synthetic **cyclic peptide** (I), in particular a AgrD-autoinducing peptide, where AgrD is a secreted **agr**-encoded peptide, where the **agr** locus controls the synthesis of virulence factor and other extracellular proteins responsible for pathogenicity in **Staphylococcus aureus**, is new.

DETAILED DESCRIPTION - A synthetic **cyclic peptide** (I) with formula (A) or $NH_2-X(n)-Z-X(y)-COOH$, and a cyclic bond between Z residue and COOH other than a thioester bond, in particular a AgrD-autoinducing peptide, where AgrD is a secreted **agr**-encoded peptide, where the **agr** locus controls the synthesis of virulence factor and other extracellular proteins responsible for pathogenicity in **Staphylococcus aureus**, is new.

X = an amino acid, its analog, peptidomimetic, or nonamide isostere;

Z' = synthetic or biosynthetic amino acid;

R = oxygen, nitrogen, sulfur or carbon;

n = 0-10; or

y = 1-10.

INDEPENDENT CLAIMS are also included for the following:

(1) a pharmaceutical composition (II) comprising (I), and a carrier;
and

(2) preparing (I).

ACTIVITY - Antibacterial.

MECHANISM OF ACTION - Inhibitor of **Agr** response. Synthetic approach was used to prepare AgrD peptides from **Staphylococcus aureus** group I and group II strains. The biological activity of the synthetic AgrD peptides was assayed using cultured **S. aureus** strains containing a beta -lactamase reporter gene fused to agrP3 promoter. This allowed activation or inhibition of the **agr** response to be monitored spectrophotometrically using a calorimetric beta -lactamase activity assay. As with their naturally derived counterparts, synthetic AgrD peptide was found to activate the **agr** response only within their own **S. aureus** class, and inhibit the **agr** response only in **S. aureus** strains from the other two classes.

USE - (I) is useful for bacterial interference, especially for treating **S. aureus** infection in a subject.
Dwg.0/3

L76 ANSWER 32 OF 34 WPIDS COPYRIGHT 2004 THOMSON DERWENT on STN
ACCESSION NUMBER: 2000-147202 [13] WPIDS
DOC. NO. CPI: C2000-046059
TITLE: New **cyclic peptides** for treating
infections with **Staphylococcus aureus**

DERWENT CLASS: A96 B04
INVENTOR(S): BEAVIS, R; JI, G; MAYVILLE, P; MUIR, T W; NOVICK, R P
PATENT ASSIGNEE(S): (UYNY) UNIV NEW YORK STATE; (UYRQ) UNIV ROCKEFELLER
COUNTRY COUNT: 87
PATENT INFORMATION:

PATENT NO KIND DATE WEEK LA PG

WO 9967286 A2 19991229 (200013)* EN 36
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW NL
OA PT SD SE SL SZ UG ZW
W: AE AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI GB
GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU
LV MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR
TT UA UG US UZ VN YU ZA ZW
AU 9947238 A 20000110 (200025)
EP 1090034 A2 20010411 (200121) EN
R: AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
JP 2002519304 W 20020702 (200246) 49

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 9967286	A2	WO 1999-US14562	19990624
AU 9947238	A	AU 1999-47238	19990624
EP 1090034	A2	EP 1999-930780	19990624
		WO 1999-US14562	19990624
JP 2002519304 W		WO 1999-US14562	19990624
		JP 2000-555937	19990624

FILING DETAILS:

PATENT NO	KIND	PATENT NO
-----	-----	-----

AU 9947238 A Based on WO 9967286
 EP 1090034 A2 Based on WO 9967286
 JP 2002519304 W Based on WO 9967286

PRIORITY APPLN. INFO: US 1998-103438 19980624

AB WO 9967286 A UPAB: 20000313

NOVELTY - **Cyclic peptides** (I) derived from AgrD peptides of **Staphylococcus aureus** are new.

DETAILED DESCRIPTION - (I) are:

(i) of formula (Ia), or
 (ii) (Ib), comprise the sequence NH₂-X_n-Z-X_y-COOH with a cyclic bond, other than thioester, between Z and COOH:

X = amino acid (aa), aa analog, peptidomimetic or non-amide isostere;

Z = synthetic or natural aa;

R = oxygen, nitrogen or carbon;

n = 0-10;

y = 1-10

INDEPENDENT CLAIMS are also included for the following:

(a) composition containing (I) and a carrier; and

(b) method for production of (I).

ACTIVITY - **Antibacterial**.

MECHANISM OF ACTION - (I) inhibits the **agr** response (associated with release of virulence factors) of **S.**

aureus. Normally an AgrD peptide is produced by **S.**

aureus that (i) activates the **agr** response in strains of a single group but (ii) interferes with this response in strains of different groups.

USE - (I) are used to treat infections by **S. aureus**

ADVANTAGE - (I), which mimic the AgrD peptides but have the thiololactone bond replaced by a different cyclic bond, lack the **agr**-inducing activity of the native peptide for specific bacterial strains, but retains the inhibitory activity against other strains. (I) can be produced in excellent yield by selective intermolecular cyclization/cleavage reaction of the resin-bound, unprotected, linear analog.

Dwg.0/3

L76 ANSWER 33 OF 34 WPIDS COPYRIGHT 2004 THOMSON DERWENT on STN
 ACCESSION NUMBER: 1999-370885 [31] WPIDS
 DOC. NO. CPI: C1999-109456
 TITLE: **Cyclic antibacterial peptides**
 for treating and preventing bacterial infection in humans and animals.
 DERWENT CLASS: B04 B07 C03 D22
 INVENTOR(S): AFFAS, Z M; BYCROFT, B W; CHAN, W C; MCDOWELL, P W;
 STEWART, G S A B; WILLIAMS, P
 PATENT ASSIGNEE(S): (UYNO-N) UNIV NOTTINGHAM
 COUNTRY COUNT: 83
 PATENT INFORMATION:

PATENT NO KIND DATE WEEK LA PG

WO 9926968 A1 19990603 (199931)* EN 32

RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW NL
 OA PT SD SE SZ UG ZW
 W: AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI GB GD
 GE GH GM HR HU ID IL IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MD
 MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT UA
 UG US UZ VN YU ZW

AU 9912490 A 19990615 (199944)

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 9926968	A1	WO 1998-GB3497	19981124
AU 9912490	A	AU 1999-12490	19981124

FILING DETAILS:

PATENT NO	KIND	PATENT NO
AU 9912490	A Based on	WO 9926968

PRIORITY APPLN. INFO: GB 1997-24859 19971126

AB WO 9926968 A UPAB: 20000516

NOVELTY - **Cyclic peptides** (I), their acid or base addition salts and prodrugs are new.

DETAILED DESCRIPTION - (I) have the formula:

A, B, E and D = residues of natural or synthetic amino acids (aa), optionally substituted;

m = 0-5;

n = 1-5; ;

X = sulfur, oxygen or NR₃;

R₁, R₂ and R₃ = hydrogen, 1-6C alkyl, acyl or alkoxy carbonyl, or, where m = 0, R₁ and R₂ may complete a heterocyclic ring;

each Z = residue as defined for A;

the peptide bonds in (I) are optionally substituted on nitrogen by 1-6C alkyl.

When n = 1, X = sulfur, and R₁ = R₂ = hydrogen and when:

(i) ABED = Asp-Phe-Ile-Met respectively, (Z)_m is not H-Tyr-Ser-Thr;

(ii) ABED = Asn-Ala-Tyr-Phe respectively, (Z)_m is not H-Asp-Ile;

(iii) ABED = Asp-Phe-Leu-Leu respectively, (Z)_m is not H-Tyr-Ile-Asn;

or

(iv) ABED = Ser-Ser-Leu-Phe respectively, (Z)_m is not

H-Gly-Val-Asn-Ala.

INDEPENDENT CLAIMS are included for the following:

(1) articles coated and/or impregnated with (I); and

(2) compounds of formula (II):

where:

(i) AlBlElDl = Asp-Phe-Ile-Met respectively, Y is H-Tyr-Ser-Thr;

(ii) AlBlElDl = Asn-Ala-Tyr-Phe respectively, Y is H-Asp-Ile;

(iii) AlBlElDl = Asp-Phe-Leu-Leu respectively, Y is H-Tyr-Ile-Asn; or

(iv) AlBlElDl = Ser-Ser-Leu-Phe respectively, Y is H-Gly-Val-Asn-Ala.

ACTIVITY - **Antibacterial.**

MECHANISM OF ACTION - (I) are modulators of the **agr** response; agonists of this response antagonize early stages of virulence, specifically production of cell wall proteins (so are preferred for prophylaxis), while antagonists interfere with the later stages of virulence, specifically secretion of exotoxin (so are preferred for treatment). **Staphylococcus aureus** KH1187A was grown in nutrient broth containing 1.92 µg/ml of the **peptide** ZA78, **cyclo**(Ac-Cys-Asp-Phe-Ile-Met), then the culture supernatant analyzed by electrophoresis and staining with Coomassie Blue. The band for toxic shock syndrome toxin-1, clearly visible in a control culture grown in absence of ZA78, was not detectable.

USE - (I) and (II) are used to treat or prevent bacterial infections in humans and animals, also for incorporation into medical articles, e.g. tampons (to prevent toxic shock syndrome), prostheses and wound dressings. Particularly they are active against staphylococci, specifically

Staphylococcus aureus but also coagulase-negative *S. epidermidis*, implicated in infection of medical implants.

ADVANTAGE - Some (I) are active against many different bacterial strains.
Dwg.0/0

L76 ANSWER 34 OF 34 WPIDS COPYRIGHT 2004 THOMSON DERWENT on STN
ACCESSION NUMBER: 1997-225834 [20] WPIDS
DOC. NO. CPI: C1997-072303
TITLE: Treatment of bacterial infections in mammals - by administering an inhibitor of a global regulator of pathogenesis genes.
DERWENT CLASS: B05
INVENTOR(S): BAO, Y; BOGGS, A; CONTAG, P R; FEDERSPIEL, N A; HEBERT, A; HECKER, S; MALOUIN, F
PATENT ASSIGNEE(S): (MICR-N) MICROCID PHARM INC
COUNTRY COUNT: 25
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 9711690	A2	19970403	(199720)*	EN	138
RW: AT BE CH DE DK ES FI FR GB GR IE IT LU MC NL PT SE					
W: AU CA CU DE IL JP MX NZ					
AU 9671686	A	19970417	(199732)		
WO 9711690	A3	19970912	(199749)		
US 6020121	A	20000201	(200013)		

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 9711690	A2	WO 1996-US15435	19960925
AU 9671686	A	AU 1996-71686	19960925
WO 9711690	A3	WO 1996-US15435	19960925
US 6020121	A Provisional	US 1995-4626P	19950929
		US 1996-672215	19960625

FILING DETAILS:

PATENT NO	KIND	PATENT NO
AU 9671686	A Based on	WO 9711690

PRIORITY APPLN. INFO: US 1996-672215 19960625; US 1995-4626P
19950929

AB WO 9711690 A UPAB: 19970516

Treatment or prophylaxis of a bacterial infection in a mammal, by admin. of a cpd. comprising an inhibitor of a global regulator (IGR) of pathogenesis genes (PG), is new. Also claimed is a method of treating a bacterial infection by administering an inhibitor of one or more pathogenesis factors modulated by *agr*, *sar*, *sae* or *xpr*. Also new is a method of screening for an IGR of PG, by determining whether a test cpd. alters the level of activity of that IGR. Also claimed is a method for making an **antibacterial** agent (ABA), comprising screening for the ABA by measuring the ability of ABA to alter the level of activity of global regulator of PG; and synthesising ABA in sufficient amount for admin. to the patient. Also claimed are pharmaceutical compsns. containing an inhibitor of a global pathogene regulator and a carrier.

USE - The cpds., termed antipathogenics, are of use in both human and veterinary infections. They inhibit the naturally occurring,

growth-dependent induction of the P3-driven RNAIII and those downstream virulence factors under its influence. Treatment with the cpds. may have any of the following effects: attenuating host inflammatory response, decreasing load of bacterial toxins, disfavours colonisation of the host, or aiding clearance of infection by the host, and potentiation of the effect of traditional drugs, either by weakening the bacteria or relocating them to a drug or host factor accessible category.

ADVANTAGE - The method represents a new approach to bacterial infections, in contrast to prior art **bactericidal**, bacteriostatic, or antibiotic cpds., which either kill the microbes or inhibit their growth. The new method targets bacterial pathogenesis, or virulence factors, reducing or eliminating expression of exoproteins, toxins, or other pathogenic factors, associated with specific genes. Even resistant strains of bacteria, partic. staphylococci, most notably **S. aureus**, which have developed resistance to methicillin or are multi-drug resistant, can be regulated by the cpds. to alter the course of an infection.

Dwg.0/20

=> file home

FILE 'HOME' ENTERED AT 17:38:01 ON 30 JAN 2004

=>

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:44:54 ; Search time 35.0735 Seconds
(without alignments)
40.730 Million cell updates/sec

Title: US-10-032-950-1

Sequence: 1 GVMXSSLF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: A_Geneseq_19Jun03:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1960.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	97.4	9	21	AA67859
2	38	97.4	9	21	AA67860
3	38	97.4	9	23	AA50907
4	38	97.4	9	23	AA51002
5	38	97.4	9	23	AA51003
6	38	97.4	9	23	AA51004
7	38	97.4	9	23	AB07160
8	37	94.9	9	21	AA67851
9	37	94.9	9	21	AA67861

ID	Score	Query Match	Length	DB ID	Description
10	37	94.9	9	23	AB53540
11	37	94.9	9	23	AA50899
12	37	94.9	9	23	AA50908
13	37	94.9	9	23	AA51005
14	37	94.9	9	23	AB07161
15	36	92.3	9	19	AA38323
16	36	92.3	9	23	AA51001
17	36	92.3	9	23	AB07159
18	36	92.3	9	24	AB84631
19	36	92.3	47	24	AB84634
20	34	87.2	9	21	AA67855
21	34	87.2	9	21	AA67856
22	34	87.2	9	23	AB53544
23	34	87.2	9	23	AB53545
24	34	87.2	9	23	AA50903
25	34	87.2	9	23	AA50904
26	33	84.6	9	21	AA67852
27	33	84.6	9	23	AB53541
28	33	84.6	9	23	AA50900
29	32	82.1	9	21	AA67857
30	32	82.1	9	23	AB53546
31	32	82.1	9	23	AA50905
32	32	82.1	895	23	AB392623
33	31	79.5	9	21	AA67854
34	31	79.5	9	23	AB53543
35	31	79.5	9	23	AA50902
36	31	79.5	243	22	AB51172
37	30	76.9	465	22	AB53547
38	30	76.9	479	21	AA50910
39	30	76.9	549	21	AA50919
40	30	76.9	560	21	AA50918
41	30	76.9	911	23	AB26657
42	30	76.9	911	23	AB26744
43	29	74.4	9	21	AA67853
44	29	74.4	9	23	AB53542
45	29	74.4	9	23	AA50901

ALIGNMENTS

RESULT 1
AA67859
ID AA67859 standard; peptide; 9 AA.
XX
XX
AC AA67859;
XX
DT 25-APR-2000 (first entry)
XX
XX
DE Staphylococcus aureus AgrDII derived peptide sequence.
XX
XX
KM Staphylococcus aureus infection; treatment; AgrD; agr response;
XX
XX
XX virulence factor.
XX
XX
OS Staphylococcus aureus.
XX
XX
PN MO9967286-A2.
XX
XX
PD 29-DEC-1999.
XX
XX
PF 24-JUN-1999; 99WO-US14562.
XX
XX
PR 24-JUN-1998; 98US-0103438.
XX
XX
PA (YRQ) UNIV ROCKEFELLER.
XX
XX (YNY) UNIV NEW YORK STATE.
XX
XX
PI Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
XX
XX
XX WPI; 2000-147202/13.
XX
XX
PT New cyclic peptides for treating infections with Staphylococcus aureus
PT

XX Examples; Page 22; 37pp; English.

PS

XX This sequence represents the Staphylococcus aureus AgrDII derived

CC peptide. The invention relates to AgrD derived cyclic peptides, a

CC composition containing a peptide and a carrier, and a method for the

CC production of the cyclic peptides. The peptide inhibits the agr response,

CC which is normally associated with the release of virulence factors of

CC Staphylococcus aureus. An AgrD peptide is produced by S. aureus that

CC activates the agr response in strains of a single group, but interferes

CC with this response in strains of different groups. The peptides and

CC composition containing them can be used to treat infections by S. aureus.

XX

SQ Sequence 9 AA;

Query Match 97.4%; Score 38; DB 21; Length 9;

Best Local Similarity 88.9%; Pred. No. 9.3e+05;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSLF 9

Db 1 GVNAXSSLF 9

RESULT 2

AA67860

ID AAY67860 standard; peptide; 9 AA.

AC AAY67860;

XX

XX 25-APR-2000 (first entry)

DT

XX Staphylococcus aureus AgrDII derived peptide sequence.

DE

XX Staphylococcus aureus infection; AgrD; agr response; treatment;

KW virulence factor.

KM

XX Staphylococcus aureus.

OS

XX WO9967286-A2.

FN

XX 29-DEC-1999.

PD

XX 24-JUN-1999; 99WO-US14562.

PF

XX 24-JUN-1998; 98US-0103438.

PR

XX (UYRQ) UNIV ROCKEFELLER.

PA (UYNY) UNIV NEW YORK STATE.

XX

PI Muir TW, Mayville P, Novick RP, Ji G, Beavis R;

PI

XX WPI; 2000-147202/13.

DR

XX New cyclic peptides for treating infections with Staphylococcus aureus

PT

XX

PS Examples; Page 22; 37pp; English.

XX This sequence represents the Staphylococcus aureus AgrDII derived

CC peptide. The invention relates to AgrD derived cyclic peptides, a

CC composition containing a peptide and a carrier, and a method for the

CC production of the cyclic peptides. The peptide inhibits the agr response,

CC which is normally associated with the release of virulence factors of

CC Staphylococcus aureus. An AgrD peptide is produced by S. aureus that

CC activates the agr response in strains of a single group, but interferes

CC with this response in strains of different groups. The peptides and

CC composition containing them can be used to treat infections by S. aureus.

XX

SQ Sequence 9 AA;

Query Match 97.4%; Score 38; DB 21; Length 9;

Best Local Similarity 88.9%; Pred. No. 9.3e+05;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSLF 9

Db 1 GVNAXSSLF 9

RESULT 3

AA50907

ID AAM50907 standard; Peptide; 9 AA.

XX

AC AAM50907;

XX

DT 08-MAY-2002 (first entry)

DE

XX Protected peptide used in cyclic peptide production.

DE

XX Staphylococcus aureus; AgrD; agr response; inhibitor; antibiotic;

KW antibacterial; infection; therapy; cyclic.

KW

XX Synthetic.

OS

XX

XX Key Location/Qualifiers

FT Modified-site 1 /note= "2-Gly"

FT Modified-site 5 /note= "Ser(tbu)"

FT Modified-site 6 /note= "Ser(Bzl)"

FT Modified-site 7 /note= "Ser(Bzl)"

FT Misc-difference 5 /note= "note linked to residue 9 to form cyclic peptide"

FT Misc-difference 9 /note= "note linked to residue 5 to form cyclic peptide"

FT

XX US6337385-B1.

PN

XX 08-JAN-2002.

PD

XX 24-JUN-1999; 99US-033511.

PF

XX 24-JUN-1998; 98US-090402P.

PR

XX (UYRQ) UNIV ROCKEFELLER.

PA (UYNY) UNIV NEW YORK STATE.

XX

PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;

PI

XX WPI; 2002-170774/22.

DR

XX Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial

PT interference and for treating Staphylococcus aureus infection in a

PT subject

XX

PS Disclosure; Column 14; 18pp; English.

XX

XX The present sequence is that of a protected peptide used in an

CC example of the preparation of novel synthetic cyclic peptides of

CC the invention (see AAM50899-906). The peptide corresponds to the

CC Staphylococcus aureus AgrDII sequence with a Cys5 to Ser mutation

CC (lactone). It was synthesised on a Wang-resin using an Fmoc

CC N-alpha protection strategy. Following chain assembly, the peptide

CC was cleaved from the support and the Ser-5 residue deprotected by

CC treatment with a trifluoroacetic acid:anisole:water mixture

CC (90:5:5) for 4 hr. The partially protected peptide-alpha

CC carboxylates were then dissolved in DMF and treated with PyBOP

CC and a catalytic amount of dimethylaminopyridine. Cyclization was

CC complete after 2 hr. The remaining protecting groups were then

CC removed by treatment with HF and the peptide purified by HPLC.

CC The cyclic peptide is capable of inhibiting the agr response of

CC Staphylococcus aureus. The thiolactone structure within native
 CC AgrD peptides is required for activation of this response.
 CC Replacement of the thiol ester component of the cyclic ring
 CC structure with a lactone (as in the present case) or a lactam can
 CC destroy agr response activating activity while preserving and
 CC enhancing inhibitory activity. The cyclic peptides are useful for
 CC bacterial interference, especially for the treatment of S. aureus
 CC infection.

CC Sequence 9 AA;

Query Match 97.4%; Score 38; DB 23; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSLF 9
 Db 1 GVNAXSSLF 9

RESULT 4
 AAMS1002
 ID AAMS1002 standard; Peptide; 9 AA.

AC AAMS1002;

DT 08-MAY-2002 (first entry)

DE AgrD2 linear thioester peptide.

KW AgrD2; agr response; inhibitor; antibiotic; antibacterial;
 infection; therapy.

OS Staphylococcus aureus.
 Synthetic.

FT Modified-site 9

FT Location/Qualifiers

FT /note="C-terminal thioester"

PN US6337385-B1.

PD 08-JAN-2002.

PF 24-JUN-1999; 99US-0339511.

PR 24-JUN-1998; 98US-090402P.

PA (UYRQ) UNIV ROCKEFELLER.

PA (UYNV) UNIV NEW YORK STATE.

PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;

PS WPI; 2002-170774/22.

PT Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial

PT interference and for treating Staphylococcus aureus infection in a

PT subject -

PS Example 1; Column 9; 18pp; English.

CC The present sequence is that of a novel synthetic AgrD2 linear
 CC thioester peptide. The peptide is derived from the cyclic AgrD2
 CC peptide of Staphylococcus aureus group II. AgrD2 is a secreted
 CC agr-encoded peptide, where the agr locus controls the synthesis of
 CC virulence factor and other extracellular proteins responsible for
 CC pathogenicity in S. aureus. The biological activity of the
 CC synthetic peptide was assayed using cultured S. aureus strains
 CC containing a beta-lactamase reporter gene fused to the agrD3
 CC promoter. This allowed activation or inhibition of the agr
 CC response to be monitored spectrophotometrically. Unlike an AgrD2
 CC thiolactone cyclic peptide (see AAMS1001), the present peptide
 CC was unable to either activate or inhibit the agr response, even

CC when added to cultured cells at 10^6 concentrations. The invention
 CC provides claimed cyclic peptides (see AAMS0839-906 and AAMS0999)
 CC and methods for preparing them. The cyclic peptides are useful for
 CC bacterial interference, especially for the treatment of S. aureus
 CC infection.

CC Sequence 9 AA;

Query Match 97.4%; Score 38; DB 23; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSLF 9
 Db 1 GVNAXSSLF 9

RESULT 5
 AAMS1003
 ID AAMS1003 standard; Peptide; 9 AA.

AC AAMS1003;

DT 08-MAY-2002 (first entry)

DE AgrD2 linear free acid peptide.

KW AgrD2; agr response; inhibitor; antibiotic; antibacterial;
 infection; therapy.

OS Staphylococcus aureus.
 Synthetic.

PN US6337385-B1.

PD 08-JAN-2002.

PF 24-JUN-1999; 99US-0339511.

PR 24-JUN-1998; 98US-090402P.

PA (UYRQ) UNIV ROCKEFELLER.

PA (UYNV) UNIV NEW YORK STATE.

PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;

PS WPI; 2002-170774/22.

PT Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial

PT interference and for treating Staphylococcus aureus infection in a

PT subject -

PS Example 1; Column 9; 18pp; English.

CC The present sequence is that of a novel synthetic AgrD2 linear
 CC free acid peptide. The peptide is derived from the cyclic AgrD2
 CC peptide of Staphylococcus aureus group II. AgrD2 is a secreted
 CC agr-encoded peptide, where the agr locus controls the synthesis of
 CC virulence factor and other extracellular proteins responsible for
 CC pathogenicity in S. aureus. The biological activity of the
 CC synthetic peptide was assayed using cultured S. aureus strains
 CC containing a beta-lactamase reporter gene fused to the agrD3
 CC promoter. This allowed activation or inhibition of the agr
 CC response to be monitored spectrophotometrically. Unlike an AgrD2
 CC thiolactone cyclic peptide (see AAMS1001), the present peptide
 CC was unable to either activate or inhibit the agr response, even
 CC when added to cultured cells at 10^6 concentrations. The invention
 CC provides claimed cyclic peptides (see AAMS0839-906 and AAMS0999)
 CC and methods for preparing them. The cyclic peptides are useful for
 CC bacterial interference, especially for the treatment of S. aureus
 CC infection.

QY	1	GVNAXSLP	9	97.4%;	Score 38;	DB 23;	Length 9;
				Best Local Similarity	88.9%;	Pred. No. 9.3e+05;	
Matches	8;	Conservative	0;	Mismatches	1;	Indels	0;
Db	1	GVNAXSSLP	9				

RESULT 6	
ID	AAM51004 standard; Peptide; 9 AA.
AC	AAM51004;
DT	08-MAY-2002 (first entry)
DE	AgrD2 lactone cyclic peptide.
XX	
XX	AgrD2; agr response; inhibitor; antibiotic; antibacterial;
XX	infection; therapy; lactone; cyclic.
XX	
OS	Staphylococcus aureus.
OS	Synthetic.
XX	
FH	Key
FT	Misc-difference 5 Location/Qualifiers
FT	Misc-difference 5 /note= "note linked to residue 9 to form cyclic peptide"
FT	
FT	Misc-difference 9 /note= "note linked to residue 5 to form cyclic peptide"
XX	
XX	US6337385-B1.
XX	
PN	08-JAN-2002.
PD	
PE	24-JUN-1999; 99US-0339511.
PR	24-JUN-1998; 98US-030402P.
XX	
XX	(UTRQ) UNIV ROCHELLE.
PA	(UYNV) UNIV NEW YORK STATE.
XX	
PI	Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX	
DR	WPI; 2002-170774/22.
XX	
PT	Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
PT	infection and for treating Staphylococcus aureus infection in a
XX	subject
PS	Example 1; Column 9; 18pp; English.
XX	

The present sequence is that of a novel synthetic AgrD2 lactone cyclic peptide in which residue 5 of the peptide is linked to residue 9 via a lactone bond. The peptide is derived from an AgrD2 peptide of *Staphylococcus aureus* group II. AgrD2 is a secreted agr-encoded peptide, where the agr locus controls the synthesis of virulence factor and other extracellular proteins responsible for pathogenicity in *S. aureus*. The biological activity of the synthetic peptide was assayed using cultured *S. aureus* strains containing a beta-lactamase reporter gene fused to the agr3 promoter. This allowed activation or inhibition of the agr response to be monitored spectrophotometrically. The lactone AgrD2 peptide inhibited the agr response of group I *S. aureus* strains without activating the agr response in group I, II or III strains. The invention provides claimed cyclic peptides (see AAM50899-906 and AAM50999) and methods for preparing them, especially peptides where the cyclic bond is a lactam or lactone bond. The cyclic peptides are useful for bacterial interference, especially for the treatment of *S. aureus* infection.

XX	Sequence	9 AA;	97.4%;	Score 38;	DB 23;	Length 9;
QY	Query Match		88.9%;	Pred. No. 9.3e+05;		
Db	Best Local Similarity	8;	Conservative	0;	Mismatches 1;	Indels 0;
	Matches	1 GVNAXSSLF 9				
		1 GVNAXSSLF 9				
Db	Query Match		97.4%;	Score 38;	DB 23;	Length 9;
	Best Local Similarity	88.9%;	Pred. No. 9.3e+05;			
	Matches	8;	Conservative	0;	Mismatches 1;	Indels 0;
		1 GVNAXSSLF 9				
		1 GVNAXSSLF 9				
Db	Query Match		97.4%;	Score 38;	DB 23;	Length 9;
	Best Local Similarity	88.9%;	Pred. No. 9.3e+05;			
	Matches	8;	Conservative	0;	Mismatches 1;	Indels 0;
		1 GVNAXSSLF 9				
		1 GVNAXSSLF 9				

```

RESULT 8
AA67851
ID AA67851 standard; peptide; 9 AA.
XX
XX
AC AA67851;
XX
XX
DT 25-APR-2000 (first entry)
XX
XX
DE S. aureus peptide #1 used for bacterial interference.
XX
XX
KM Staphylococcus aureus infection; cyclic peptide; AgrD; agr response;
XX
XX
OS virulence factor; treatment.
XX
XX
OS Staphylococcus aureus.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 5 /label= Unknown
XX
XX
PN WO967286-A2.
XX
XX
PD 29-DEC-1999.
XX
XX
PF 24-JUN-1999; 99WO-US14562.
XX
XX
PR 24-JUN-1998; 98US-0103438.
XX
XX
PA (UVRQ ) UNIV ROCKEFELLER.
XX
XX
PA (UVRQ ) UNIV NEW YORK STATE.
XX
XX
PI Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
XX
XX
DR WPI; 2000-147202/13.
XX
XX
PT New cyclic peptides for treating infections with Staphylococcus aureus
XX
XX
PS Claim 9; Page 26; 37pp; English.
XX
XX
CC This sequence represents a cyclic peptide derived from the Staphylococcus
CC aureus AgrD peptide. The invention relates to AgrD derived peptides,
CC a composition containing a peptide and a carrier, and a method for the
CC production of the cyclic peptides. The peptide inhibits the agr response,
CC which is normally associated with the release of virulence factors of
CC Staphylococcus aureus. An AgrD peptide is produced by S. aureus that
CC activates the agr response in strains of a single group, but interferes
CC with this response in strains of different groups. The peptides and
CC composition containing them can be used to treat infections by S. aureus.
XX
XX
SQ Sequence 9 AA;

Query Match 94.9%; Score 37; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 1 GVNAXSSLF 9
DB 1 GVNAXSSLF 9

RESULT 9
AA67861
ID AA67861 standard; peptide; 9 AA.
XX
XX
AC AA67861;
XX
XX
DT 25-APR-2000 (first entry)
XX
XX
DE Staphylococcus aureus AgrDII derived peptide sequence.
XX
XX
KM Staphylococcus aureus infection; AgrD; agr response; treatment;
XX

```

```

XX
XX
KM virulence factor.
XX
XX
OS Staphylococcus aureus.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 5 /label= Unknown
XX
XX
PN WO967286-A2.
XX
XX
PD 29-DEC-1999.
XX
XX
PF 24-JUN-1999; 99WO-US14562.
XX
XX
PR 24-JUN-1998; 98US-0103438.
XX
XX
PA (UVRQ ) UNIV ROCKEFELLER.
XX
XX
PA (UVRQ ) UNIV NEW YORK STATE.
XX
XX
PI Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
XX
XX
DR WPI; 2000-147202/13.
XX
XX
PT New cyclic peptides for treating infections with Staphylococcus aureus
XX
XX
PS Examples; Page 22; 37pp; English.
XX
XX
CC This sequence represents the Staphylococcus aureus AgrDII derived
CC peptide. The invention relates to AgrD derived cyclic peptides, a
CC composition containing a peptide and a carrier, and a method for the
CC production of the cyclic peptides. The peptide inhibits the agr response,
CC which is normally associated with the release of virulence factors of
CC Staphylococcus aureus. An AgrD peptide is produced by S. aureus that
CC activates the agr response in strains of a single group, but interferes
CC with this response in strains of different groups. The peptides and
CC composition containing them can be used to treat infections by S. aureus.
XX
XX
SQ Sequence 9 AA;

Query Match 94.9%; Score 37; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 1 GVNAXSSLF 9
DB 1 GVNAXSSLF 9

RESULT 10
ABP53540
ID ABP53540 standard; peptide; 9 AA.
XX
XX
AC ABP53540;
XX
XX
DT 13-DEC-2002 (first entry)
XX
XX
DE Cyclic peptide SEQ ID NO:1.
XX
XX
KM Cyclic peptide; Staphylococcus aureus; infection; antibacterial;
XX
XX
KM agr response inhibitor.
XX
XX
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 5 /note= "any amino acid"
XX
XX
PN US2002077453-A1.
XX
XX
PD 20-JUN-2002.
XX
XX
PF 27-DEC-2001; 2001US-0032950.
XX

```

```

XX 24-JUN-1998; 98US-090402P.
PR 24-JUN-1999; 99US-0339511.
XX (MUIR/) MUIR T W.
XX (MAYV/) MAYVILLE P.
XX (NOVI/) NOVICK R P.
XX (BEAV/) BEAVIS R.
XX (JICG/) JI G.
XX Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX WPI; 2002-681366/73.
XX New cyclic peptides, useful for treating Staphylococcus aureus
XX infections.
XX Claim 9; Page 10; 18pp; English.
XX ABP53540 to ABP53547 represent cyclic peptides (1) from the present
XX invention. The present invention also describes a method for treating
XX Staphylococcus aureus infection comprising the administration of a
XX composition comprising (i), (1) has antibacterial activity, and can be
XX used as an agr gene response inhibitor. The peptides are useful for
XX treating S. aureus infections.
XX Sequence 9 AA;
SQ
Query Match 94.9%; Score 37; DB 23; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GVNAXSSLF 9
DB 1 GVNAXSSLF 9
RESULT 11
AAMS0899
ID AAMS0899 standard; Peptide; 9 AA.
XX
XX AAMS0899;
AC
XX
XX 08-MAY-2002 (first entry)
DT
XX
XX AgrD-autoinducing cyclic peptide, inhibitor of agr response.
DE
XX
XX Staphylococcus aureus; AgrD; agr response; inhibitor; antibiotic;
XX antibacterial; infection; therapy; cyclic.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
XX FT Misc-difference 5 /note= "any amino acid"
XX FT Misc-difference 5 /note= "note linked to residue 9 to form cyclic
XX FT peptide"
XX FT Misc-difference 9 /note= "note linked to residue 5 to form cyclic
XX FT peptide"
XX US6337385-B1.
XX
XX 08-JAN-2002.
XX
XX 24-JUN-1999; 99US-0339511.
XX
XX 24-JUN-1998; 98US-090402P.
XX
XX (UYRO ) UNIV ROCKEFELLER.
XX (UYNY ) UNIV NEW YORK STATE.
XX

```

```

PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX WPI; 2002-170774/22.
XX
XX Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
XX PT interference and for treating Staphylococcus aureus infection in a
XX PT subject.
XX Claim 7; Column 19; 18pp; English.
XX
XX The present sequence is that of a novel synthetic cyclic peptide
XX of the invention that is capable of inhibiting the agr response of
XX Staphylococcus aureus. It is an AgrD-autoinducing peptide, where
XX AgrD is a secreted agr-encoded peptide and where the agr locus
XX controls the synthesis of virulence factor and other extracellular
XX proteins responsible for pathogenicity in S. aureus. Preferred
XX peptides may have the sequence NH2-X(n)-Z-X(y)-COOH, with a
XX cyclic bond between the Z residue and COOH other than a thioester
XX bond, where X is an amino acid, an amino acid analogue, a
XX peptidomimetic or non-amide isostere, Z is a synthetic or a
XX biosynthetic amino acid, n is 0-10 and y is 1-10. The cyclic bond
XX is especially a lactam or lactone bond. The thiololactone
XX structure within native AgrD peptides is required for activation
XX of the agr response. Elimination of the thiol ester component of
XX the cyclic ring structure can destroy agr response activating. A
XX claimed method of preparing a cyclic peptide involves: assembling
XX a linear peptide chain on to a solid phase resin support;
XX deprotecting the resulting protected peptide; treating the
XX deprotected peptide with neutral buffer for a time sufficient to
XX form the cyclic peptide and cleave the peptide from the support;
XX and recovering the cyclic peptide. The peptide is useful for
XX bacterial interference, especially for the treatment of S. aureus
XX infection.
XX Sequence 9 AA;
SQ
Query Match 94.9%; Score 37; DB 23; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GVNAXSSLF 9
DB 1 GVNAXSSLF 9
RESULT 12
AAMS0908
ID AAMS0908 standard; Peptide; 9 AA.
XX
XX AAMS0908;
AC
XX
XX 08-MAY-2002 (first entry)
DT
XX
XX Protected peptide used in cyclic peptide production.
DE
XX
XX Staphylococcus aureus; AgrD; agr response; inhibitor; antibiotic;
XX antibacterial; infection; therapy; cyclic.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
XX FT Modified-site 1 /note= "Z-Gly"
XX FT Modified-site 5 /label= Dpr(Boc)
XX FT Modified-site 6 /note= "Ser(Bzl)"
XX FT Modified-site 7 /note= "Ser(Bzl)"
XX FT Misc-difference 5 /note= "note linked to residue 9 to form cyclic
XX FT peptide"

```

```

FT      Masc-difference 9
FT      /note= "note linked to residue 5 to form cyclic
FT      peptide"
XX      US6337385-B1.
XX      08-JAN-2002.
XX      24-JUN-1999; 99US-0339511.
XX      24-JUN-1998; 98US-090402P.
XX      (UVRQ ) UNIV ROCKEFELLER.
XX      (UVRN ) UNIV NEW YORK STATE.
XX      Blair TW, Mayville P, Novick RP, Beavis R, Ji G;
XX      WPI; 2002-170774/22.
XX      Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
XX      interference and for treating Staphylococcus aureus infection in a
XX      subject
XX      Disclosure, Column 14; 18pp; English.
XX      The present sequence is that of a protected peptide used in an
XX      example of the preparation of novel synthetic cyclic peptides of
XX      the invention (see AAMS0899-906). The peptide corresponds to the
XX      Staphylococcus aureus AgrDII sequence with a Cys5 to diaminopropionic
XX      acid (Dpr) mutation (lactam). It was synthesized on a Wang-resin
XX      using an Fmoc N-alpha protection strategy. Following chain assembly,
XX      the peptide was cleaved from the support and the Dpr-5 residue
XX      deprotected by treatment with a trifluoroacetic acid:anisole:water
XX      mixture (90:5:5) for 4 hr. The partially protected peptide-alpha
XX      carboxylates were then dissolved in DMF and treated with PyBOP.
XX      Cyclization was complete after 2 hr. The remaining protecting groups
XX      were removed by treatment with HF and the peptide purified by HPLC.
XX      The cyclic peptide is capable of inhibiting the agr response of
XX      Staphylococcus aureus. The thiololactone structure within native
XX      AgrD peptides is required for activation of this response.
XX      Replacement of the thiol ester component of the cyclic ring
XX      structure with a lactam (as in the present case) or a lactone can
XX      destroy agr response activating activity while preserving and
XX      enhancing inhibitory activity. The cyclic peptides are useful for
XX      bacterial interference, especially for the treatment of S. aureus
XX      infection.
XX      Sequence 9 AA;
SQ
Query Match 94.9%; Score 37; DB 23; Length 9;
Best Local Similarity 100.0%, Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
CY      1 GVNAKXSLF 9
      |||||
Db      1 GVNAKXSLF 9
RESULT 13
AAMS1005
AAMS1005 standard; Peptide; 9 AA.
XX      AAMS1005;
XX      08-MAY-2002 (first entry)
XX      AgrD2 lactam cyclic peptide.
XX      AgrD2; agr response; inhibitor; antibiotic; antibacterial;
XX      infection; therapy; cyclic.
XX      Staphylococcus aureus.
XX      Synthetic.

```

XX	Key	Location/Qualifiers
FT	Misc-difference 5	/note= "any amino acid"
FT	Misc-difference 5	/note= "note linked to residue 9 to form cyclic peptide"
FT	Misc-difference 9	/note= "note linked to residue 5 to form cyclic peptide"
FT		
PN	US6337385-B1.	
XX	08-JAN-2002.	
PD		
XX	24-JUN-1999;	99US-0339511.
PF		
XX	24-JUN-1998;	98US-090402P.
PR		
XX	(UYRQ) UNIT ROCKEFELLER.	
PA	(UYNY) UNIT NEW YORK STATE.	
PA		
XX	Muir TW, Mayville P, Novick RP, Beavis R, Ji G;	
PI		
DR	WPI; 2002-170774/22.	
XX		
PS	Example 1; Column 11; 18pp; English.	
XX		
CC	The present sequence is that of a novel synthetic AgrD2 lactam	
CC	cyclic peptide in which residue 5 of the peptide is linked to	
CC	residue 9 via a lactam bond. The peptide is derived from an	
CC	AgrD2 peptide of <i>Staphylococcus aureus</i> group II. AgrD2 is a	
CC	secreted agr-encoded peptide, where the agr locus controls the	
CC	synthesis of virulence factor and other extracellular proteins	
CC	responsible for pathogenicity in <i>S. aureus</i> . The biological S.	
CC	activity of the synthetic peptide was assayed using cultured <i>S.</i>	
CC	aureus strains containing a beta-lactamase reporter gene fused to	
CC	the agr promoter. This allowed activation or inhibition of the	
CC	agr response to be monitored spectrophotometrically. The AgrD3	
CC	lactam peptide inhibited the agr response in group I <i>S. aureus</i>	
CC	strains and did not activate the agr response in group I, II or	
CC	III strains. The invention provides claimed cyclic peptides (see	
CC	CA AAM50899-906 and AAM50999) and methods for preparing them,	
CC	especially where the cyclic bond is a lactam or lactone bond.	
CC	The cyclic peptides are useful for bacterial interference,	
CC	especially for the treatment of <i>S. aureus</i> infection.	
XX		
SO	Sequence 9 AA;	
	Query Match 94.9%; Score 37; DB 23; Length 9;	
	Best Local Similarity 100.0%; Pred. No. 9.3e+05;	
	Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
GY	1 GVMAXSLF 9	
DB	1 GVMAXSLF 9	
RESULT 14		
ID	ABB07161 standard; peptide; 9 AA.	
AC	ABB07161;	
DT	13-MAR-2002 (First entry)	
DE	Peptide-mediated quorum sensing inhibitor peptide cyclo-XIII.	
XX		
XX	Autoinducer-2; AI-2; antibiotic; antibacterial; dermatological;	

KM vulnerable; pheromone; agr system; accessory gene regulator; cyclic.
 XX Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Modified-site 5 /note="diaminopropionic acid residue"
 FT
 PN W0200185664-A2.
 PD 15-NOV-2001.
 XX
 PF 10-MAY-2001; 2001WO-US15221.
 XX
 PR 10-MAY-2000; 2000US-203000P.
 PR 07-DEC-2000; 2000US-254398P.
 XX
 PA (UTPR-) UNIV PRINCETON.
 PA (QUOR-) QUOREX PHARM INC.
 PA (UYTE-) UNIV TECHNOLOGIES INT INC.
 PI Baseler BL, Dammel CS, Schauder S, Shokat K, Stein J, Surette MG;
 XX
 DR WPI; 2002-075235/10.
 XX
 PT Use of autoinducer-2 agonists or antagonists for regulating activity of
 PT autoinducer-2 receptor, regulating bacterial growth and pathogenesis,
 PT also antibiotic compositions -
 XX
 PS Disclosure; Page 33; 134pp; English.
 XX
 CC The invention relates to the use of autoinducer-2 (AI-2) agonists or
 CC antagonists for regulating activity of autoinducer-2 receptor, regulating
 CC bacterial growth and pathogenesis. Synergistic antibiotic compositions
 CC comprising inhibitors of the quorum-sensing pathway of a microorganism
 CC are also provided. Methods using such AI-2 analogues are useful for
 CC treating pathogen-associated disease states. The compounds and antibiotic
 CC compositions can be used to inhibit bacterial cell growth and/or biofilm
 CC formation on a medical device, particularly for promoting growth of skin
 CC graft replacements used in the treatment of burns and ulcers. They may
 CC also be used to aid wound repair, and to inhibit bacterial cell growth
 CC and biofilm formation in or on products or devices used for personal
 CC hygiene. The present sequence represents a inhibitor of peptide-mediated
 CC quorum sensing.
 CC
 XX
 SQ Sequence 9 AA;
 Query Match 94.9%; Score 37; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVNAXSSLF 9
 DB 1 GVNAXSSLF 9
 RESULT 15
 AAM38323
 ID AAM38323 standard; peptide; 9 AA.
 XX
 AC AAM38323;
 XX
 DT 28-MAY-1998 (first entry)
 XX
 DE Transcription inhibitory peptide 3.
 XX
 KW Transcription inhibitory peptide; inhibition; antibiotic; abscesses;
 KW endocarditis; pneumonia; osteomyelitis; virulence factor; agr regulator;
 KW agr-rnaii gene; colonisation.
 XX
 OS Synthetic.
 OS Staphylococcus aureus.
 XX

FH Key Location/Qualifiers
 FT Misc-difference 5 /note="linked to the carboxy terminus via a cyclic
 FT thioester bond"
 FT
 PN W09744349-A1.
 PD 27-NOV-1997.
 XX
 PF 22-MAY-1997; 97WO-US08791.
 XX
 PR 22-MAY-1996; 96US-0651226.
 XX
 PA (UYNY) UNIV NEW YORK STATE.
 PA Beavis R, Novick RP;
 PI
 XX
 DR WPI; 1998-018426/02.
 XX
 PT Peptide inhibiting agr-rnaii transcription in Staphylococcus aureus
 PT - useful to treat S. aureus infection
 XX
 PS Claim 7; Page 15; 24pp; English.
 XX
 CC This novel peptide sequence (or it's analogue) inhibits agr-rnaii
 CC transcription in Staphylococcus aureus (S. aureus), and is 9 amino acids
 CC long. The peptide can be added to pharmaceutical compositions to treat
 CC or prevent infections or diseases caused by S. aureus in animals, and
 CC especially in humans. They can also be combined with traditional
 CC antibiotics. S. aureus causes conditions ranging from abscesses (boils
 CC and furuncles) to life-threatening infections e.g. endocarditis,
 CC pneumonia, osteomyelitis, by secreting injurious proteins, or virulence
 CC factors. The synthesis of virulence factors is controlled by the agr
 CC regulator, which activates transcription of the genes encoding the
 CC produce RNAIII, which induces transcription of the genes encoding the
 CC virulence factors. The peptides inhibit agr-rnaii transcription in and
 CC therefore virulence factor expression. The peptides can also be used
 CC in vitro to prevent Staphylococcus aureus colonisation.
 CC
 XX
 SQ Sequence 9 AA;
 Query Match 92.3%; Score 36; DB 19; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GVNAXSSLF 9
 DB 1 GVNACSSLF 9
 Search completed: January 29, 2004, 14:52:10
 Job time : 35.0735 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:46:49 ; Search time 11.1176 Seconds

(without alignments)
77,851 Million cell updates/sec

Title: US-10-032-950-1

Perfect score: 39
Sequence: 1 GVNAXSLF 9

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 9616882 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

PIR_76:*
1: p1r:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	92.3	47	2 C89995	AgRD protein (limpo
2	35	89.7	382	2 A28067	lysosomal membrane
3	35	89.7	405	2 A60534	P2B/LAMP-1 precursor
4	35	89.7	407	2 A30200	120K lysosomal mem
5	34	87.2	333	1 C69612	ferrichrome ABC tr
6	32	82.1	895	2 T45786	receptor-protein k
7	31	79.5	389	1 S36653	probable protein - Esc
8	31	79.5	394	2 E81286	phosphotransferase
9	31	79.5	631	2 D69750	SEB24 protein - Ye
10	31	79.5	926	2 S48463	hypothetical prote
11	30	76.9	109	2 A99227	hypothetical prote
12	30	76.9	420	2 AD0105	probable sugar-pin
13	30	76.9	550	2 T02404	probable beta-gluc
14	30	76.9	633	2 T27215	hypothetical prote
15	29	74.4	269	2 S58439	transcription fact
16	29	74.4	369	2 B43715	protein M homolog
17	29	74.4	382	2 B65160	mannitol-1-phospha
18	29	74.4	382	2 AH0976	mannitol-1-phospha
19	29	74.4	398	2 F85056	hypothetical prote
20	29	74.4	416	1 A31959	ubiquitinome-associat
21	29	74.4	453	1 ZPBOC2	zip protein precu
22	29	74.4	500	1 EPPF	probable exported
23	29	74.4	618	2 AB0186	protein Y32FeA.3
24	29	74.4	655	2 G89189	hypothetical prote
25	29	74.4	746	2 T01536	hypothetical prote
26	29	74.4	963	2 T26022	hydroxymethylgluta
27	29	74.4	1054	2 A30239	hypothetical prote
28	29	74.4	1435	2 T32930	hypothetical prote
29	28	71.8	53	2 AF2557	hypothetical prote

30	28	71.8	271	2 F90350	hypothetical prote
31	28	71.8	346	2 S53491	probable galactosy
32	28	71.8	349	2 S74439	iron(III) diclrat
33	28	71.8	350	2 E64499	hypothetical prote
34	28	71.8	381	1 A53224	ubiquitinol-cytochro
35	28	71.8	381	2 S12023	ubiquitinol-cytochro
36	28	71.8	421	2 C70038	maltoase/maltodextr
37	28	71.8	424	2 C83902	maltoase/maltodextr
38	28	71.8	442	2 AF2539	manganese transpor
39	28	71.8	468	2 T24523	hypothetical prote
40	28	71.8	476	2 D86306	similar to Cytochr
41	28	71.8	480	2 C56744	hypothetical prote
42	28	71.8	514	2 T47837	beta-glucosidase-1
43	28	71.8	549	2 F64640	conserved hypothe
44	28	71.8	558	2 F82236	conserved hypothe
45	28	71.8	573	2 F81313	peptidase (M3 fami

ALIGNMENTS

RESULT 1
C89995
AgRD protein [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C/Accession: C89995
R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Qui, L.; O
ma, A.; Mizutani-Uli, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shibata, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89758; MUID:2131352; PMID:11418146
A/Accession: C89995
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-47 <KOR>
A/Cross-references: GB:BA000018; PTD:G13701831; PIDN:BA843124.1; GSPDB:GN00149
A/Experimental source: strain N315
C/Genetics:
A/Gene: agrD

Query Match 92.3%; Score 36; DB 2; Length 47;
Best Local Similarity 88.9%; Pred. No. 0.11;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSLF 9
Db 24 GVNAXSLF 32

RESULT 2
A28067
Lysosomal membrane glycoprotein LAMP-1 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 20-Aug-1999
C/Accession: A28067
R/Chen, J.W.; Cha, Y.; Yuksei, K.U.; Gracy, R.W.; August, J.T.
J. Biol. Chem. 263, 8754-8758, 1988
A/Title: Isolation and sequencing of a cDNA clone encoding lysosomal membrane glycopr
A/Reference number: A28067; MUID:88243732; PMID:3379044
A/Accession: A28067
A/Molecule type: mRNA
A/Residues: 1-382 <CHB>
A/Cross-references: GB:J01881; NID:G198706; PIDN:AAA39411.1; PID:G293692
A/Note: the authors translated the codon ATT for residue 1 as Leu and CCG for residue
C/Superfamily: lysosome-associated membrane protein
C/Keywords: glycoprotein; membrane protein

Query Match 89.7%; Score 35; DB 2; Length 382;
Best Local Similarity 77.8%; Pred. No. 2;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAKSSSLF 9
DB 256 GVNAKSSSLF 264

RESULT 3

A:Accession: A60534
A:Species: Mus musculus (house mouse)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 03-May-1996
C:Accession: A60534
R:Hefferman, M.; Yousefi, S.; Dennis, J.W.
A:Title: Molecular characterization of P2B/LAMP-1, a major protein target of a metacystic
A:Reference number: A60534; PMID:9002989; PMID:2676155
A:Accession: A60534
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-405 <HEP>
C:Superfamily: lysosome-associated membrane protein

Query Match 89.7%; Score 35; DB 2; Length 405;
Best Local Similarity 77.8%; Pred. No. 2.1;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAKSSSLF 9
DB 279 GVNAKSSSLF 287

RESULT 4

A:Accession: A30200
A:Species: Rattus norvegicus (Norway rat)
C:Date: 18-Oct-1989 #sequence_revision 30-Sep-1991 #text_change 20-Aug-1999
C:Accession: A30200; S03331
R:Howe, C.L.; Granger, B.L.; Hull, M.; Green, S.A.; Gabel, C.A.; Helenius, A.; Mellman, P.; Natl. Acad. Sci. U.S.A. 85, 7577-7581, 1988
A:Title: Derived protein sequence, oligosaccharides, and membrane insertion of the 120-k glycoproteins.
A:Reference number: A30200; PMID:89017240; PMID:3174652
A:Accession: A30200
A:Molecule type: mRNA
A:Residues: 1-407 <HOW>
A:Cross-references: EMBL:J01672
A:Note: The authors translated the codon GGG for residue 15 as Val
R:Himeno, M.; Noguchi, Y.; Sasaki, H.; Tanaka, Y.; Furuno, K.; Kono, A.; Sasaki, Y.; Kat
FBS Lett. 244, 351-356, 1989
A:Title: Isolation and sequencing of a cDNA clone encoding 107 kDa sialoglycoprotein in
A:Reference number: S03331; PMID:9153580; PMID:2920835
A:Accession: S03331
A:Molecule type: mRNA
A:Residues: 22-407 <HM>
A:Cross-references: EMBL:X14765; NID:G56577; PIDN:CAA32873.1; PID:G56578
A:Note: Part of this sequence, including the amino end of the mature protein, was confir
C:Superfamily: lysosome-associated membrane protein
C:Keywords: glycoprotein; membrane protein
F:22-407/Product: lysosomal membrane glycoprotein, 107K #status experimental <MAT>

Query Match 89.7%; Score 35; DB 2; Length 407;
Best Local Similarity 77.8%; Pred. No. 2.1;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAKSSSLF 9
DB 281 GVNAKSSSLF 289

RESULT 5
C69812
ferrichrome ABC transporter (permease) homolog yfmd - Bacillus subtilis
C:Species: Bacillus subtilis

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: C69812

R:Kunst, P.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C: Bron, S.; Brouillet, S.; Bruchet, C.V.; Caldwell, B.; Caprano, V.; Carter, N.M.;
A: Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gal
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M
Koeter, P.; Konigstein, G.; Krogan, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portere
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl
A:Authors: Schleich, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akouch, M.; Tamakoshi, A.; Tanaka, T.; Teperera, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winers, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A:Reference number: A69580; PMID:98044033; PMID:9384377
A:Accession: C69812

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-333 <KUN>
A:Cross-references: GB:299108; GB:AL009126; NID:G2633055; PIDN:CAE12580.1; PID:G26330
A:Experimental source: strain 168
C:Genetics:

A:Gene: yfmd
C:Superfamily: ferrichrome ABC transporter

Query Match 87.2%; Score 34; DB 1; Length 333;
Best Local Similarity 77.8%; Pred. No. 2.9;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAKSSSLF 9
DB 99 GVNAKSSSLF 107

RESULT 6

A:Accession: T45786
A:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Mar-2000
C:Accession: T45786
R:Dejany, M.; Berger, C.; Cocke, R.; Grellier, F.; Laidie, M.; Mewes, H.W.; Lemcke, K
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23013
A:Accession: T45786
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-895
A:Cross-references: EMBL:AL133452
A:Experimental source: cultivar Columbia; BAC clone F26013
C:Genetics:
A:Map position: 3
A:Note: F26013.190
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom

Query Match 82.1%; Score 32; DB 2; Length 895;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAKSSSLF 9
DB 122 GVNAKSSSLF 130

RESULT 7
S36653
kpsd protein - Escherichia coli
C:Species: Escherichia coli
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Jul-1999
C:Accession: S36653; E48492; C42644

R, Roberts, I.S.
submitted to the EMBL Data Library, August 1993
C/Species: Campylobacter jejuni
A/Reference number: S36649
A/Accession: S36653
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1389 <RDB>
A/Cross-references: EMBL:X74567; NID:G397404; PIDN:CA52659.1; PID:G397409
R, Pazanli, C.; Rosenow, C.; Boulinols, G.J.; Bronner, D.; Jann, K.; Roberts, I.S.
J. Bacteriol. 175, 5978-5983, 1993
A/Title: Molecular analysis of region 1 of the Escherichia coli K5 antigen gene cluster:
A/Reference number: A48492; MUID:93388530; PMID:8397187
A/Accession: E48492
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 22369 <PAZ>
A/Cross-references: GB:X74567
R, Steenbergen, S.M.; Wrona, T.J.; Vint, E.R.
J. Bacteriol. 174, 1099-1108, 1992
A/Title: Functional analysis of the sialyltransferase complexes in Escherichia coli K1
A/Reference number: A42644; MUID:92138601; PMID:1735705
A/Accession: C42644
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 106-125, 'D', 127-244, 'N', 246-297, 'S', 299-300, 'VI', 303-379, 'W', 381-385, 'W', 387
A/Cross-references: GB:W6370; NID:G146947; PIDN:AAA24214.1; PID:G146950
A/Experimental source: K1, strain EVL
A/Note: sequence extracted from NCBI backbone (NCBI:79370, NCBI:79381)
C/Genetics:
A/Gene: kpsD
C/Superfamily: kpsD protein

Query Match 79.5%; Score 31; DB 1; Length 389;
Best Local Similarity 87.5%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GVNASSL 8
Db 127 GVNASSL 134

RESULT 8
E81286
probable polysaccharide modification protein Cj1413c [imported] - Campylobacter jejuni
C/Species: Campylobacter jejuni
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #ext_change 03-Jun-2002
C/Accession: E81286
R, Packhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chittin
C.W.; Quail, M.; Rajandream, N.A.; Rutherford, K.M.; Vanlille, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A/Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A/Reference number: A81250; MUID:20150912; PMID:10688204
A/Accession: E81286
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-994 <PAR>
A/Cross-references: GB:AL139078; GB:AL111168; NID:G6968723; PIDN:CAB73837.1; PID:G696884
A/Experimental source: serotype O2, strain NCTC 11168
C/Genetics:
A/Gene: Cj1413c
C/Superfamily: kpsD protein

Query Match 79.5%; Score 31; DB 2; Length 394;
Best Local Similarity 87.5%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GVNASSL 8
Db 126 GVNASSL 133

RESULT 9
D69750

phosphotransferase system enzyme II homolog ybfS - Bacillus subtilis
C/Species: Bacillus subtilis
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #ext_change 20-Jun-2000
C/Accession: D69750
R, Kunst, F.; Ogasawara, N.; Mosser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Buschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fuma, S.; Gallizi, A.; Gal
lechi, J.; Harwood, C.R.; Hensel, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A/Reference number: A63580; MUID:98044033; PMID:9384377
A/Accession: D69750
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-631 <KUN>
A/Cross-references: GB:Z99105; GB:AL009126; NID:G2632457; PIDN:CAB12029.1; PID:G26325
A/Experimental source: strain 168
C/Genetics:
A/Gene: ybfS
C/Superfamily: phosphotransferase system N-acetylglucosamine-specific enzyme II; phos
nzyme II, factor III homology

Query Match 79.5%; Score 31; DB 2; Length 631;
Best Local Similarity 75.0%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GVNASSL 8
Db 172 GVNASSL 179

RESULT 10
S48463
SEC24 protein - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein YIL109C
C/Species: Saccharomyces cerevisiae
C/Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #ext_change 19-Apr-2002
C/Accession: S48463
R, Bowman, S.; Churcher, C.
submitted to the EMBL data Library, September 1994
A/Reference number: S48463
A/Accession: S48463
A/Molecule type: DNA
A/Residues: 1-926 <BOW>
A/Cross-references: GB:Z47047; EMBL:Z38125; NID:G603997; PID:G763237; CSPDB:GN00009;
C/Genetics:
A/Gene: SCD:SEC24; SEC24; MIPS:YIL109C
A/Cross-references: MIPS:YIL109C; SCD:50001371
A/Map position: 9L
C/Function:
A/Description: involved in endoplasmic reticulum to Golgi transport; required for vesi

Query Match 79.5%; Score 31; DB 2; Length 926;
Best Local Similarity 75.0%; Pred. No. 48;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VNASSLF 9
Db 787 INATSLF 794

RESULT 11
A99227
hypothetical protein S800775 [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus
 C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
 C:Accession: A99227
 R:Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyer, M.J.; Chan-
 Jongs, I.; Jettles, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Hoc, H.P.; Redder, R.
 arett, R.A.; Ragan, M.A.; Sersen, C.W.; Van der Oost, J.
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A94420; PMID:20083487; PMID:10617197
 A:Accession: B94879
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-109 <KUR>
 A:Cross-references: GB:AE006641; NID:G13813950; PIDN:AAK41072.1; GSPDB:GN00155
 C:Genetics:
 A:Gene: SS00775

Query Match 76.9%; Score 30; DB 2; Length 109;
 Best Local Similarity 55.6%; Pred. No. 8.1;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVNAXSLF 9
 |||||
 Db 99 GINAKVKSIF 107

RESULT 12
 AD0105
 probable sugar-binding protein YPO0856 [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
 C:Accession: AD0105
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; PMID:21470413; PMID:11586360
 A:Accession: AD0105
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-420 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC89703.1; PID:G15978930; GSPDB:GN00175
 C:Genetics:
 A:Gene: YPO0856
 C:Superfamily: maltose-binding protein

Query Match 76.9%; Score 30; DB 2; Length 420;
 Best Local Similarity 66.7%; Pred. No. 35;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVNAXSLF 9
 |||||
 Db 248 GINAKVKSIF 256

RESULT 13
 T02404
 probable beta-glucosidase homolog F411.30 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
 C:Accession: T02404; B94879
 R:Rounsley, S.D.; Jin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul,
 submitted to the EMBL Data Library, May 1998
 A:Description: Arabidopsis thaliana chromosome II BAC F411 genomic sequence.
 A:Reference number: Z14667
 A:Accession: T02404
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-560 <ROU>
 A:Cross-references: EMBL:AC004521; NID:G3128166; PIDN:AACT6095.1; PID:G3128191
 A:Experimental source: cultivar Columbia
 R:Jin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Varadhan, S.E.; Umayam, L.; Tallon,
 euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A94420; PMID:20083487; PMID:10617197
 A:Accession: B94879
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-560 <STO>
 A:Cross-references: GB:AE002093; NID:G3128191; PIDN:AACT6095.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2G44490; F411.30
 A:Map position: 2
 A:Insertions: 3/2; 33/3; 76/3; 102/1; 157/2; 242/3; 281/2; 354/1; 400/1; 438/2
 C:Superfamily: Agrobacterium beta-glucosidase

Query Match 76.9%; Score 30; DB 2; Length 560;
 Best Local Similarity 66.7%; Pred. No. 48;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVNAXSLF 9
 |||||
 Db 322 GINAYSLF 330

RESULT 14
 T27215
 hypothetical protein Y57G11C.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
 C:Accession: T27215
 R:McMurray, A.
 submitted to the EMBL Data Library, September 1997
 A:Reference number: Z20330
 A:Accession: T27215
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-633 <ML>
 A:Cross-references: EMBL:Z29281; PIDN:CAB16503.1; GSPDB:GN00022; CESP:Y57G11C.1
 A:Experimental source: clone Y57G11C
 C:Genetics:
 A:Gene: CESP:Y57G11C.1
 A:Map position: 4
 A:Insertions: 70/1; 329/3
 C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

Query Match 76.9%; Score 30; DB 2; Length 633;
 Best Local Similarity 75.0%; Pred. No. 55;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 VNAKSLF 9
 |||||
 Db 328 INAFSLF 335

RESULT 15
 S58439
 transcription factor AKR - chicken
 N:Alternate names: avian knotted-related protein; homeotic protein AKR
 C:Species: Gallus gallus (chicken)
 C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 24-Sep-1999
 C:Accession: S58439; S58440
 R:Ryan, A.K.; Tejeda, M.L.; May, D.L.; Dubrava, M.; Deeley, R.G.
 submitted to the EMBL Data Library, April 1995
 A:Description: Isolation and characterization of the chicken homeodomain protein AKR.
 A:Reference number: S58439
 A:Accession: S58439
 A:Molecule type: mRNA
 A:Residues: 1-269 <RYA>
 A:Cross-references: EMBL:U25353; NID:G857681; PIDN:AAA83567.1; PID:G857682
 R:Ryan, A.K.; Tejeda, M.L.; May, D.L.; Dubrava, M.; Deeley, R.G.
 Nucleic Acids Res. 23, 3252-3259, 1995
 A:Title: Isolation and characterization of the chicken homeodomain protein AKR.

A;Reference number: S58440; MUID:95396587; PMID:7667102
 A;Accession: S58440
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 35-94 <RT>
 A;Cross-references: EMBL:U25353
 C;Function:
 A;Description: negative regulator of apolipoprotein (major egg yolk apolipoprotein)
 A;Note: expressed in liver during early embryogenesis and in non-hepatic adult tissues
 C;Superfamily: unassigned homeobox proteins; homeobox homology
 C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;36-95/Domain: homeobox homology <HOX>

Query Match 74.4%; Score 29; DB 2; Length 269;
 Best Local Similarity 66.7%; Pred. No. 38;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GYNAKSLF 9
 |||||
 Db 222 GANAGSLF 230

Search completed: January 29, 2004, 14:58:19
 Job time : 12.1176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen, Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:45:44 ; Search time 5.82353 Seconds
(without alignments)
72.678 Million cell updates/sec

Title: US-10-032-950-1
Perfect score: 39
Sequence: 1 GUNAXSLP 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	35	89.7	406 1	LMP1_MOUSE
2	35	89.7	407 1	LMP1_RAT
3	31	79.5	389 1	KSS5_ECOLI
4	31	79.5	407 1	LMP1_CRIGR
5	31	79.5	631 1	YBFS_BACSU
6	31	79.5	926 1	SC24_YEAST
7	29	74.4	269 1	AKR_CHICK
8	29	74.4	369 1	MX_STRPY
9	29	74.4	382 1	MTLD_ECOL6
10	29	74.4	382 1	MTLD_ECOLI
11	29	74.4	382 1	MTLD_ECOLI
12	29	74.4	382 1	MTLD_ECOLI
13	29	74.4	382 1	MTLD_ECOLI
14	29	74.4	411 1	LCVB_SYMP7
15	29	74.4	416 1	LMP1_HUMAN
16	29	74.4	453 1	UCR2_BOVIN
17	29	74.4	488 1	UZIP_DROME
18	29	74.4	1054 1	HMD1_YEAST
19	28	71.8	117 1	RBFA_LEBIN
20	28	71.8	230 1	TRPF_SYNEL
21	28	71.8	346 1	YAOD_SCHPO
22	28	71.8	381 1	CYB_CHLRE
23	28	71.8	391 1	CYB_CHUSM
24	28	71.8	442 1	WNTH_ANASP
25	28	71.8	468 1	YJ95_CABEL
26	28	71.8	558 1	YB45_VIBCH
27	28	71.8	694 1	PPK_CAMJE
28	28	71.8	1200 1	DP3A_CAMJE
29	28	69.2	106 1	YAHN_ECOLI
30	27	69.2	102 1	Y014_TREPA
31	27	69.2	225 1	NARI_ECOLI
32	27	69.2	237 1	YM9A_YEAST
33	27	69.2	260 1	TATD_ECOLI

34	27	69.2	293 1	IF35_ARATH	004202 arabidopsis
35	27	69.2	325 1	YJ40_YEAST	P47120 saccharomyc
36	27	69.2	379 1	Y528_SYNY3	055518 synechocyst
37	27	69.2	395 1	NUSA_HELPU	094166 heliobacte
38	27	69.2	395 1	NUSA_HELPU	P55977 heliobacte
39	27	69.2	492 1	GLPK_AQUHE	066746 aquifex aeo
40	27	69.2	533 1	GPRT_ANASP	084712 arabidops
41	27	69.2	571 1	URE1_YEREN	P31494 yersinia en
42	27	69.2	571 1	URE1_YEREP	092419 yersinia pe
43	27	69.2	571 1	URE1_YERPS	P52313 yersinia ps
44	27	69.2	756 1	MAC1_SCHPO	010268 schizosacch
45	27	69.2	1024 1	Y075_MYCCE	P47321 mycoplasma

ALIGNMENTS

RESULT 1
LMP1_MOUSE
ID LMP1_MOUSE STANDARD: PRT: 406 AA.
AC P1438; Q62020;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Lysosome-associated membrane glycoprotein 1 precursor (LAMP-1) (LGP-A)
DE (LGP-120) (CD107A) (P2B).
GN LAMP1 OR LAMP-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
[1]
PP SEQUENCE FROM N.A.
RX MEDLINE=90307738; PubMed=2142158;
RA Granger B.L., Green S.A., Gabel C.A., Howe C.L., Wellman I.,
RA Helentius A.;
RT "Characterization and cloning of Igpl10, a lysosomal membrane
RT glycoprotein from mouse and rat cells."
RT J. Biol. Chem. 265:12036-12043(1990).
[2]
PP SEQUENCE FROM N.A.
RX Helentius A., Yusefi S., Dennis J.W.;
RA Submitted (FEB-1990) to the EMBL/Genbank/DBJ databases.
[3]
PP SEQUENCE OF 25-406 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=88243732; PubMed=3379044;
RA Chen J.W., Cha Y., Yuksel K.U., Gracy R.W., August J.T.;
RT "Isolation and sequencing of a cDNA clone encoding lysosomal membrane
RT glycoprotein mouse LAMP-1. Sequence similarity to proteins bearing
RT onco-differentiation antigens."
RT J. Biol. Chem. 263:8754-8758(1988).
[4]
PP DISULFIDE BONDS.
RX MEDLINE=90237040; PubMed=2332434;
RA Atreburn L.M., Barles B.J., August J.T.;
RT "The disulfide structure of mouse lysosome-associated membrane
RT protein 1."
RT J. Biol. Chem. 265:7419-7423(1990).
CC -1- FUNCTION: PRESENTS CARBOHYDRATE LIGANDS TO SELECTINS. ALSO
CC IMPLICATED IN TUMOR CELL METASTASIS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Lysosomal.
CC THIS PROTEIN SHUTTLES BETWEEN LYSOSOMES, ENDOPLASM, AND THE PLASMA
CC MEMBRANE.
CC -1- PTM: O- AND N-GLYCOSYLATED. SOME OF THE N-GLYCANS ATTACHED TO
CC LAMP-1 ARE POLYLACTOSAMINOGLYCANS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LAMP FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)


```

FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 293 294 GV -> EF (IN REF. 3).
FT CONFLICT 329 329 N -> T (IN REF. 3).
FT CONFLICT 356 357 SD -> VT (IN REF. 3).
SQ SEQUENCE 407 AA; 43969 MW; 25947490749ATC68 CRC64;

Query Match 89.7%; Score 35; DB 1; Length 407;
Best Local Similarity 77.8%; Pred. No. 1.2;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GVNAXSSLF 9
Db 281 GVNAXSSLF 285

RESULT 3
KSS5_ECOLI STANDARD; PRT; 389 AA.
AC P42216;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Capsule polysaccharide export protein kps5.
GN KPS5.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X5;
RX MEDLINE=93388530; PubMed=8397187;
RA Pazzani C., Rosenow C., Boulnois G.J., Bromner D., Jann K.,
RA Roberts I.S.;
RT "Molecular analysis of region 1 of the Escherichia coli K5 antigen
RT gene cluster: a region encoding proteins involved in cell surface
RT expression of capsular polysaccharide.";
RL J. Bacteriol. 175:5978-5983(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X74567; CA52659.1; -
CC PIR; S36653; S36653.
CC Pfam; PF05159; Capsule_synth.1.
CC Polysaccharide transport; Transport.
SQ SEQUENCE 389 AA; 46381 MW; 73058122C28027DE CRC64;

Query Match 79.5%; Score 31; DB 1; Length 389;
Best Local Similarity 87.5%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GVNAXSSL 8
Db 127 GVNAXSSL 134

RESULT 4
LMP1_CRIGR STANDARD; PRT; 407 AA.
AC P49129;

```

```

DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Lysosome-associated membrane glycoprotein 1 precursor (LAMP-1)
DE (lysosomal membrane glycoprotein A) (Lgp A).
GN LAMP1 OR LgpA.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97021428; PubMed=8667788;
RA Uchayakumar S., Granger B.L.;
RT "Cell surface accumulation of overexpressed hamster lysosomal
RT membrane glycoproteins.";
RL Cell. Mol. Biol. Res. 41:405-420(1995).
CC -1- FUNCTION: PRESENTS CARBOHYDRATE LIGANDS TO SELECTING. ALSO
CC -1- IMPLICATED IN TUMOR CELL METASTASIS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein. Lysosomal.
CC THIS PROTEIN SHUTTLES BETWEEN LYSOSOMES, ENDOSOMES, AND THE PLASMA
CC MEMBRANE (BY SIMILARITY).
CC -1- PTM: O- AND N-GLYCOSYLATED; SOME OF THE N-GLYCANS ATTACHED TO
CC LAMP-1 ARE POLYLACTOSAMINOGLYCANS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LAMP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L18986; AAC37682.1; -
CC InterPro; IPR002000; Lamp.
DR Pfam; PF01299; Lamp; 1.
DR PRINTS; PRO0336; LYSASOCTDMP.
DR PROSITE; PS00310; LAMP_1; 2.
DR PROSITE; PS00311; LAMP_2; 1.
KW Transmembrane; Glycoprotein;
FT SIGNAL 1 21 Lysosome; Signal.
FT CHAIN 22 407 BY SIMILARITY.
FT 1. LYSOSOME-ASSOCIATED MEMBRANE GLYCOPROTEIN
FT 1. LUMENAL (POTENTIAL).
FT DOMAIN 22 371 POTENTIAL.
FT TRANSMEM 372 395 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 396 407 FIRST LUMENAL DOMAIN.
FT DOMAIN 422 189 HINGE.
FT DOMAIN 190 219 SECOND LUMENAL DOMAIN.
FT DISULFID 220 371 BY SIMILARITY.
FT DISULFID 36 75 BY SIMILARITY.
FT DISULFID 150 186 BY SIMILARITY.
FT DISULFID 223 260 BY SIMILARITY.
FT DISULFID 328 365 BY SIMILARITY.
FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 297 297 N-LINKED (GLCNAC...) (POTENTIAL)
FT CARBOHYD 304 304 N-LINKED (GLCNAC...) (POTENTIAL)
FT CARBOHYD 312 312 N-LINKED (GLCNAC...) (POTENTIAL)
SO SEQUENCE 407 AA; 43786 MM; 651002040F86B3D CRC64;

Query Match 79.5%; Score 31; DB 1; Length 407;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

CY 1 GVNAXSSSLF 9
Db 281 GVNAXSSSLF 289

RESULT 5
YBFS_BACSU STANDARD; PRT; 631 AA.
ID P39816;
AC 01-FEB-1995 (Rel. 31, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative PTS system IIBC component ybfs (EC 2.7.1.69).
GN YBFS.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OK NCBI_taxid=1423;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Hagi K., Liu H., Yasumoto K., Takahashi H., Yoshikawa H.;
RT "Sequence analysis of the 70kb region between 17 and 23 degree of the
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=9804403; PubMed=9384377;
RA Kunze F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bortner M.G., Bessieres P., Bolotin A., Borchert S.,
RA Bortner R., Bortner L., Brans A., Brann M., Brignell S.C., Bron S.,
RA Broutier S., Brusch C.V., Caldwell B., Capuano V., Carter N.W.,
RA Choi S.K., Codani U.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S., Emerson P.T.,
RA Ertan C.D., Ertan K.M., Fabbre C., Ferrari E., Follmer D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallazzi A., Galleron N.,
RA Chim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy B.U., Hagi K., Hahne U., Hatwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mestl D., Nakai S., Noback M.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogawa A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Priescan E., Pujic P., Putetelle D., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadleir Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Taccanti E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpestra F., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassatoli A.,
RA Viati A., Wambut R., Wedler E., Wedler H., Wettengesser T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumsstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis".
RL Nature 390:249-256(1997).
RN (3)
RP SEQUENCE OF 515-631 FROM N.A.
RC STRAIN=168 / 6GM;
RX MEDLINE=95270606; PubMed=7751298;
RA Tolner B., Ubbink-Kox T., Poolman B., Konings W.N.;

```

```

RT "Characterization of the proton/glutamate symport protein of Bacillus
RT subtilis and its functional expression in Escherichia coli.";
RL J. Bacteriol. 177:2863-2869(1995).
CC -1- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT
CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE
CC -TRANSPORT SYSTEM. THE IICD DOMAINS CONTAIN THE SUGAR BINDING SITE
CC AND THE TRANSMEMBRANE CHANNEL. THE IIA DOMAIN CONTAINS THE PRIMARY
CC PHOSPHORYLATION SITE (THE DONOR IS PHOSPHO-HPR); IIA TRANSFERS ITS
CC PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO
CC THE SUGAR (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Protein N-phosphotransferase + sugar = protein
CC histidine + sugar phosphate.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Contains 1 PTS EIIa domain.
CC -1- SIMILARITY: Contains 1 PTS EIIb domain.
CC -1- SIMILARITY: Contains 1 PTS EIIc domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL; AB006424; BA33132.1; -
DR EMBL; Z99105; CAB12028.1; -
DR EMBL; U15147; AAA82877.1; -
DR PIR; D69750; D69750.
DR HSSP; P20166; IGPR.
DR Subtilisin; BG11014; YBFS.
DR InterPro; IPR001127; PTS_EIIA.
DR InterPro; IPR001986; PTS_EIIB.
DR InterPro; IPR003382; PTS_EIIC.
DR Pfam; PF00358; PTS_EIIA_1; 1.
DR Pfam; PF00367; PTS_EIIB; 1.
DR Pfam; PF02378; PTS_EIIC; 1.
DR ProDom; PD002243; PTS_EIIA; 1.
DR ProDom; PD004476; PTS_EIIB; 1.
DR TIGRPFAMs; TIGR00826; EIIb_glc; 1.
DR TIGRPFAMs; TIGR00830; PTM; 1.
DR PROSITE; PS00371; PTS_EIIA_1; 1.
DR PROSITE; PS00371; PTS_EIIB_CVS; 1.
DR Hypothetical protein; Phosphotransferase system; Sugar transport;
KW Transferrase; Phosphorylation; Transmembrane; Complete proteome.
FT MOD_RES 203 203 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 419 419 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 567 567 PHOSPHORYLATION (BY SIMILARITY).
FT DOMAIN 1 477 EIIc domain.
FT DOMAIN 478 631 EIIb domain.
FT DOMAIN 632 777 EIIa domain.
FT TRANSMEM 12 32 POTENTIAL.
FT TRANSMEM 56 76 POTENTIAL.
FT TRANSMEM 77 97 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 149 169 POTENTIAL.
FT TRANSMEM 173 193 POTENTIAL.
FT TRANSMEM 196 216 POTENTIAL.
FT TRANSMEM 243 263 POTENTIAL.
FT TRANSMEM 274 294 POTENTIAL.
FT TRANSMEM 298 316 POTENTIAL.
FT TRANSMEM 350 370 POTENTIAL.
SQ SEQUENCE 631 AA; 66144 MM; 88388A72E8C85B9B CRC64;

Query Match 79.5%; Score 31; DB 1; Length 631;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

CY 1 GVNAXSSSL 8
Db 172 GVNAXSSSL 179

```

```

RESULT 6
SC24_YEAST
ID SC24_YEAST STANDARD; PRT; 926 AA.
AC P40482;
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Protein transport protein Sec24 (Abnormal nuclear morphology 1).
GN SEC24 OR ANU1 OR YIL109C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
CX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX PubMed=9169870;
RA Churcher C.M., Bowden S., Baddock K., Bankier A., Brown D.,
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harteis D.B., Horsnell T., Hunt S., Jagels K., Jones M., Lye G.,
RA Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Bartell B.G.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
RA Nature 387:84-87 (1997).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=20219212; PubMed=10753972;
RA Peng R., De Antoni A., Gallwitz D.;
RT "Evidence for overlapping and distinct functions in protein transport
RT of coat protein Sec24p family members.";
RL J. Biol. Chem. 275:11521-11528 (2000).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=20177547; PubMed=10712514;
RA Kurihara T., Hamamoto S., Gimeno R.E., Kaiser C.A., Schekman R.,
RA Yoshinaka T.;
RT "Sec24p and Issip function interchangeably in transport vesicle
RT formation from the endoplasmic reticulum in Saccharomyces
RT cerevisiae."
RL Mol. Biol. Cell 11:983-998 (2000).
RN [4]
RP FUNCTION: COMPONENT OF THE COPII COAT THAT COVERS ER-DERIVED
RN VESICLES INVOLVED IN TRANSPORT FROM THE ENDOPLASMIC RETICULUM TO
RN THE GOLGI APPARATUS. COPII ACTS IN THE CYTOPLASM TO PROMOTE THE
RN TRANSPORT OF SECRETORY, PLASMA MEMBRANE, AND VACUOLAR PROTEINS
RN FROM THE ENDOPLASMIC RETICULUM TO THE GOLGI COMPLEX.
RN -!- SUBUNIT: COPII IS COMPOSED OF AT LEAST FIVE PROTEINS: THE SEC23/24
RN COMPLEX, THE SEC13/31 COMPLEX AND SAR1.
RN -!- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
RN -!- SIMILARITY: BELONGS TO THE SEC23/SEC24 FAMILY. SEC24 SUBFAMILY.
RN -----
RN This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See http://www.isb-sib.ch/announce/
RN or send an email to license@isb-sib.ch).
RN -----
DR EMBL; Z38125; CAAB6271.1; -
DR PIR; S48463; S48463
DR PDB; 1M2V; 20-SEP-02
DR SGD; S0001371; SEC24.
DR GO; GO:0005515; F:protein binding activity; IDA.
DR GO; GO:0006914; P:autophagy; IMP.
DR InterPro; IPR001974; Gelsolin.
DR InterPro; IPR006900; Sec23 helical.
DR InterPro; IPR006896; Sec23_trunk.
DR InterPro; IPR006895; ZF-Sec23_Sec24.
DR Pfam; PF00626; Gelsolin; 1.
DR Pfam; PF04815; Sec23_helical; 1.
DR Pfam; PF04811; Sec23_trunk; 1.
DR Pfam; PF04810; ZF-Sec23_Sec24; 1.
DR Transport; Protein transport; Golgi stack; Endoplasmic reticulum;
KW

```

```

KW Multigene family; 3D-structure.
FT DOMAIN 231 256 ZINC FINGER-LIKE.
FT DOMAIN 110 114 POLY-GLN.
FT DOMAIN 157 160 POLY-PRO.
SQ SEQUENCE 926 AA; 103635 MW; 35E2BD4CC75899 CRC64;
Query Match
Best Local Similarity 79.5%; Score 31; DB 1; Length 926;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
DY 2 VNAXSLF 9
DB 787 INRTSLF 794

```

```

RESULT 7
AKR_CHICK
ID AKR_CHICK STANDARD; PRT; 269 AA.
AC Q90555;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeodomain protein AKR (Avian knotted-related protein).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
CX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95396587; PubMed=7667102;
RA Ryan A.K., Tejeda M.L., May D.L., Dubrava M., Deeley R.G.;
RT "Isolation and characterization of the chicken homeodomain protein
RT AKR."
RL Nucleic Acids Res. 23:3252-3259 (1995).
CC -!- FUNCTION: BINDS TO THE 5' ELEMENT OF THE APOVDLII GENE AND
CC REPRESS ITS TRANSCRIPTION.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE TALE/TGIF HOMEBOX FAMILY.
CC -!- SIMILARITY: Contains 1 homeobox domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U25353; AAA83567.1; -
DR PIR; S58439; S58439.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SW00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; FALSE_NEG.
DR PROSITE; PS50071; HOMEBOX 2; 1.
KW Transcription regulation; Repressor; Homeobox; DNA-binding;
KW Nuclear Protein.
FT DNA_BIND 35 97 HOMEBOX (TLF-TYPE).
SQ SEQUENCE 269 AA; 29442 MW; 107435DDCC0C2253 CRC64;
Query Match
Best Local Similarity 74.4%; Score 29; DB 1; Length 269;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
DY 1 GVNAXSLF 9
DB 222 GANAQSLF 230

```

RESULT 8

```

MX STRPY
ID MX STRPY STANDARD; PRT; 369 AA.
AC P16946;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Virulence factor-related M protein precursor.
GN ENNX.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=serotype M49;
RX MEDLINE=90078078; PubMed=2687231;
RA Haanes E.U., Cleary P.P.;
RT "Identification of a divergent M protein gene and an M protein-related
  gene family in Streptococcus pyogenes serotype 49."
RL U. Bacteriol. 172:6397-6408(1989).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
  an amide bond (potential).
CC -1- SIMILARITY: BELONGS TO THE M PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M23689; AA26919.1; -
DR EMBL; M31790; AA26875.1; -
DR PIR; B43715; B43715.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR006192; LPxTG.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02370; M7.
DR Pfam; PF04650; YsIRK_signal; 1.
DR PRINTS; PR00015; GPOSANCHOR.
DR TIGRFAMs; TIGR01167; LPxTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YsIRK_signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Virulence: Phagocytosis; Cell wall; Peptidoglycan-anchor; Repeat;
  Antigen; Coiled coil; Signal.
FT SIGNAL 1 41
FT CHAIN 1 339 VIRULENCE FACTOR-RELATED M PROTEIN.
FT PROPEP 340 369 REMOVED BY SORTASE (POTENTIAL).
FT DOMAIN 129 200 2 X REPEATS, TYPE A.
FT REPEAT 129 159 A-1.
FT REPEAT 171 200 A-2.
FT DOMAIN 132 241 3 X REPEATS, TYPE B.
FT REPEAT 132 241 B-1.
FT REPEAT 132 148 B-1.
FT REPEAT 173 190 B-2.
FT REPEAT 224 241 B-3.
FT DOMAIN 150 215 3 X REPEATS, TYPE C.
FT REPEAT 150 160 C-1.
FT REPEAT 192 201 C-2.
FT REPEAT 206 215 C-3.
FT DOMAIN 302 335 GLY/PRO-RICH (CELL WALL-SPANNING).
FT SITE 336 340 LPxTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 339 339 AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ SEQUENCE 369 AA; 40639 MW; 4AD6D5A72F637839 CRC64;

```

```

Query Match      74.4%; Score 29; DB 1; Length 369;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GVNAXSSL 8
|||||

```

```

Db 43 GVNATTSLS 50
RESULT 9
MTLD_ECOLI6 STANDARD; PRT; 382 AA.
ID MTLD_ECOLI6
AC Q8FCE7;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Mannitol-1-phosphate 5-dehydrogenase (EC 1.1.1.17).
GN MTLD OR C4417.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=06:H1 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
  Raeko D., Buckles E.L., Liu S.-R., Boutin A., Hackett J., Stroud D.,
  Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
  Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
  of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -1- CATALYTIC ACTIVITY: D-mannitol 1-phosphate + NAD(+) = D-fructose
  6-phosphate + NADH.
CC -1- SIMILARITY: Belongs to the mannitol dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE016768; AAN82853.1; ALT_INT.
DR HAMAP; MF_00196; -1.
DR InterPro; IPR000669; Mannitol_dh.
DR Pfam; PF01232; Mannitol_dh; 1.
DR PRINTS; PR00084; MTLDHDEHGNASE.
DR PROSITE; PS00974; MANNITOL_DHGENASE; 1.
KW Oxidoreductase; NAD; Complete proteome.
FT NP_BIND 3 14 NAD (BY SIMILARITY).
FT SEQUENCE 382 AA; 41124 MW; 3121F924FE9357B7 CRC64;

```

```

Query Match      74.4%; Score 29; DB 1; Length 382;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GVNAXSSL 8
|||||
Db 62 GVNAXSSI 69
RESULT 10
MTLD_ECOLI STANDARD; PRT; 382 AA.
ID MTLD_ECOLI
AC P03424;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Mannitol-1-phosphate 5-dehydrogenase (EC 1.1.1.17).
GN MTLD OR B3600.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=362;
RN [1]
RP SEQUENCE FROM N.A.

```



```

RC STRAIN=K12;
RX MEDLINE=86288055; PubMed=3135464;
RA Davis T., Yamada M., Elgort M., Sailer M.H. Jr.;
RT "Nucleotide sequence of the mannitol (mtl) operon in Escherichia
RT coli.";
RN Mol. Microbiol. 2:405-412(1988).
RP SEQUENCE FROM N.A., AND REVISIONS.
RC STRAIN=K12;
RX MEDLINE=9186837; PubMed=1954486;
RA Uaiang W., Wu L.F., Tomlich J., Sailer M.H. Jr., Nitsch W.G.;
RT "Corrected sequence of the mannitol (mtl) operon in Escherichia
RT coli.";
RL Mol. Microbiol. 4:2003-2006(1990).
RN SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=94316500; PubMed=8041620;
RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
RT region from 76.0 to 81.5 minutes."
RL Nucleic Acids Res. 22:2576-2586(1994).
RN SEQUENCE OF 369-382 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=9411964; PubMed=8300537;
RA Figge R.M., Ramsesler T.M., Sailer M.H. Jr.;
RT "The mannitol repressor (MtlR) of Escherichia coli."
RL J. Bacteriol. 176:840-847(1994).
RN SEQUENCE OF 1-25.
RX MEDLINE=85006766; PubMed=6384188;
RA Novorny M.J., Reizer J., Esch F., Sailer M.H. Jr.;
RT "Purification and properties of D-mannitol-1-phosphate dehydrogenase
RT and D-glucitol-6-phosphate dehydrogenase from Escherichia coli."
RL J. Bacteriol. 159:988-990(1984).
CC -1- CATALYTIC ACTIVITY: D-mannitol 1-phosphate + NAD(+) = D-fructose
CC 6-phosphate + NADH.
CC -1- SUBUNIT: Monomer.
CC -1- SIMILARITY: Belongs to the mannitol dehydrogenase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X51359; CA35744.1; -
DR EMBL; U00039; AAB18577.1; -
DR EMBL; AE000438; AAC76624.1; -
DR EMBL; X06794; CA29954.1; ALT_SEQ.
DR EMBL; U03845; AAA2661.1; -
DR PIR; B65160; B65160.
DR EcoGene; EG10616; mtlD.
DR HAMAP; MF_00196; -; 1.
DR InterPro; IPR000669; Mannitol_dh.
DR Pfam; PF01232; Mannitol_dh; 1.
DR PRINTS; PRO0084; MTLDHNGNSE.
DR PROSITE; PS00974; MANNITOL_DHGENSE; 1.
KM Oxidoreductase; NAD: Complete proteome.
FT NP_BIND 3 14 NAD (BY SIMILARITY).
FT CONFLICT 86 86 A -> R (IN REF. 1 AND 2).
SQ SEQUENCE 382 AA; 41139 MW; 1AC4028C150A7B2 CRC64;

Query Match 74.4%; Score 29; DB 1; Length 382;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GYNAVSSI 8
DB 62 GYNAVSSI 69

```

```

RESULT 11
MTLD_KLEPN STANDARD; PRT; 382 AA.
ID MTLD_KLEPN
AC O9XEM6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Mannitol-1-phosphate 5-dehydrogenase (EC 1.1.1.17).
GN MTLD.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=573;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=103-SP14 / KAY2026;
RX MEDLINE=21100248; PubMed=1164312;
RA Otle S., Lengeler J.W.;
RT "The mtl genes and the mannitol-1-phosphate dehydrogenase from
RT Klebsiella pneumoniae KAY2026."
RL FEMS Microbiol. Lett. 194:221-227(2001).
CC -1- CATALYTIC ACTIVITY: D-mannitol 1-phosphate + NAD(+) = D-fructose
CC 6-phosphate + NADH.
CC -1- SIMILARITY: Belongs to the mannitol dehydrogenase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF16095; AAD5386.1; -
DR HAMAP; MF_00196; -; 1.
DR InterPro; IPR000669; Mannitol_dh.
DR Pfam; PF01232; Mannitol_dh; 1.
DR PRINTS; PRO0084; MTLDHNGNSE.
DR PROSITE; PS00974; MANNITOL_DHGENSE; 1.
KM Oxidoreductase; NAD.
FT NP_BIND 3 14 NAD (BY SIMILARITY).
SQ SEQUENCE 382 AA; 41130 MW; D9927B6C06C5344 CRC64;

Query Match 74.4%; Score 29; DB 1; Length 382;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GYNAVSSI 8
DB 62 GYNAVSSI 69

RESULT 12
MTLD_SALTI STANDARD; PRT; 382 AA.
ID MTLD_SALTI
AC O8Z2E0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Mannitol-1-phosphate 5-dehydrogenase (EC 1.1.1.17).
GN MTLD OR STY4110 OR T3833.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill V., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina W.,

```

RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jørgensen K.,
 RA Krogsh A., Larsen T.S., Lechner S., Mølle S., O'Gaora P., Perry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrall B.G.,
 RT "Complete genome sequence of a multiple drug resistant *Salmonella*
 RT enterica serovar Typhimurium CT18.",
 RL Nature 413:848-852(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.,
 RT "Comparative genomics of *Salmonella enterica* serovar Typhimurium Ty2
 RT and CT18.",
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -1- CATALYTIC ACTIVITY: D-mannitol 1-phosphate + NAD(+) = D-fructose
 CC -1- 6-phosphate + NADH.
 CC -1- SIMILARITY: Belongs to the mannitol dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AL627280; CAD03308.1; -
 DR EMBL: AE016847; AA071314.1; -
 DR HAMAP: MF_00196; -; 1.
 DR InterPro: IPR000669; Mannitol_dh.
 DR Pfam: PF01232; Mannitol_dh; 1.
 DR PRINTS: PR00084; MTLDDHGNASE.
 DR PROSITE: PS00974; MANNITOL_DHCENASE; 1.
 KW Oxidoreductase; NAD; Complete proteome.
 FT NP BIND 3 14 NAD (BY SIMILARITY).
 SQ SEQUENCE 382 AA; 40869 MW; C562C574161443B1 CRC64;
 QY Query Match 74.4%; Score 29; DB 1; Length 382;
 DB Best Local Similarity 75.0%; Pred. No. 30;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GVNAXSSL 8
 DB 62 GVNAXSSI 69
 RESULT 13
 ID MTLTID_SALTY STANDARD; PRT; 382 AA.
 AC 08ZLE7; 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Mannitol-1-phosphate 5-dehydrogenase (EC 1.1.1.17).
 GN MTLTID OR STM3686.
 OS *Salmonella typhimurium*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Salmonella*.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hsu S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.,
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium

RT LT2.",
 RL Nature 413:852-856(2001).
 CC -1- CATALYTIC ACTIVITY: D-mannitol 1-phosphate + NAD(+) = D-fructose
 CC -1- 6-phosphate + NADH.
 CC -1- SIMILARITY: Belongs to the mannitol dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE008871; AA122545.1; -
 DR STYGEN: SG2727; mtlD.
 DR HAMAP: MF_00196; -; 1.
 DR InterPro: IPR000669; Mannitol_dh.
 DR Pfam: PF01232; Mannitol_dh; 1.
 DR PRINTS: PR00084; MTLDDHGNASE.
 DR PROSITE: PS00974; MANNITOL_DHCENASE; 1.
 KW Oxidoreductase; NAD; Complete proteome.
 FT NP BIND 3 14 NAD (BY SIMILARITY).
 SQ SEQUENCE 382 AA; 40899 MW; C5781F86C507E1 CRC64;
 QY Query Match 74.4%; Score 29; DB 1; Length 382;
 DB Best Local Similarity 75.0%; Pred. No. 30;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GVNAXSSL 8
 DB 62 GVNAXSSI 69
 RESULT 14
 ID LCVY SYN7 STANDARD; PRT; 411 AA.
 AC 055276; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lycopene beta cyclase (EC 1.14.-.-).
 GN CRT1 OR LCVY.
 OS *Synechococcus* sp. (strain PCC 7942) (Anacystis nidulans R2).
 OC Cyanobacteria; Chroococcales; *Synechococcus*.
 OX NCBI_TaxID=1140;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=95003701; PubMed=7919981;
 RA Cunningham F.X. Jr., Sun Z., Chamovitz D., Hirschberg J., Gantt E.,
 RT "Molecular structure and enzymatic function of lycopene cyclase from
 RT the cyanobacterium *Synechococcus* sp strain PCC7942.",
 RL Plant Cell 6:1107-1121(1994).
 CC -1- FUNCTION: CATALYZES THE DOUBLE CYCLIZATION REACTION WHICH CONVERTS
 CC LYCOPENE TO BETA-CAROTENE AND NEUROSPORINE TO BETA-ZEAXANTHENE.
 CC -1- ENZYME REGULATION: INHIBITED BY THE BLEACHING HERBICIDE 2-(4-
 CC METHYLPHENOXY)TRIFLUORAMINE HYDROCHLORIDE (WPTA).
 CC -1- PATHWAY: Carotenoid biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X74599; CA52677.1; -
 DR InterPro: IPR000205; NAD_binding.
 DR InterPro: IPR000103; Pyridine_redox_2.
 DR PRINTS: PR00469; PNDPDTASEII.
 KW Oxidoreductase; NAD; Carotenoid biosynthesis.

```

FT NP_BIND 4 32 NAD (POTENTIAL).
SQ SEQUENCE 411 AA: 46085 MW: C46CGB28557A2 CRC64;

Query Match 74.4%; Score 29; DB 1; Length 411;
Best Local Similarity 75.0%; Pct. No. 32;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GYVAXSSL 8
Db 299 GINAXSSL 306

RESULT 15
LMP1_HUMAN
ID LMP1_HUMAN STANDARD; PRT; 416 AA.
AC P11279;
DT 01-JUN-1989 (Rel. 11, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lysosome-associated membrane glycoprotein 1 precursor (LAMP-1)
DE (CD107a antigen).
GN LAMP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_Taxid:9606;
RX P11279
RA Fukuoka M., Vitale J., Matreson J., Carlsson S.R.;
RT "Cloning of cDNAs encoding human lysosomal membrane glycoproteins, h-
RT lamp-1 and h-lamp-2. Comparison of their deduced amino acid
RT sequences";
RT J. Biol. Chem. 263:18920-18928(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=93252065; PubMed=8517882;
RA Savada R., Jardine K.A., Fukuoka M.;
RT "The genes of major lysosomal membrane glycoproteins, lamp-1 and
RT lamp-2. 5'-flanking sequence of lamp-2 gene and comparison of exon
RT organization in two genes";
RT J. Biol. Chem. 268:9014-9022(1993).
RN [3]
RP SEQUENCE OF 32-416 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=88234502; PubMed=331762;
RA Vitale J., Carlsson S.R., Siebert P.D., Fukuoka M.;
RT "Molecular cloning of cDNAs encoding lamp A, a human lysosomal
RT membrane glycoprotein with apparent Mr approximately equal to
RT 120,000";
RT Proc. Natl. Acad. Sci. U.S.A. 85:3743-3747(1988).
RN [4]
RP SEQUENCE OF 28-58.
RX MEDLINE=89104438; PubMed=2912382;
RA Mane S.M., Marzella L., Bailton D.F., Holt V.K., Cha Y.,
RA Hildreth J.E.K., August J.T.;
RT "Purification and characterization of human lysosomal membrane
RT glycoproteins";
RT Arch. Biochem. Biophys. 268:360-378(1989).
RN [5]
RP DISULFIDE BONDS.
RX MEDLINE=90062189; PubMed=2584229;
RA Carlsson S.R., Fukuoka M.;
RT "Structure of human lysosomal membrane glycoprotein 1. Assignment of
RT disulfide bonds and visualization of its domain arrangement";
RT J. Biol. Chem. 264:20526-20531(1989).
RN [6]
RP POLYLACTOSAMINOGLYCANS.
RX MEDLINE=91056099; PubMed=2243102;
RA Carlsson S.R., Fukuoka M.;
RT "The polylactosaminoglycans of human lysosomal membrane glycoproteins
RT lamp-1 and lamp-2. Localization on the peptide backbones";
RT J. Biol. Chem. 265:20488-20495(1990).

```

[illegible]

```
FT CARBOHYD 210 210 O-LINKED (GALNAC. . .)
FT CARBOHYD 222 222 N-LINKED (GLCNAC. . .)
FT CARBOHYD 227 227 (POLYLACTOSAMINOGLYCAN).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .)
FT CARBOHYD 240 240 (POLYLACTOSAMINOGLYCAN).
FT CARBOHYD 248 240 N-LINKED (GLCNAC. . .)
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .)
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .)
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .)
FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .)
FT CARBOHYD 379 379 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 32 39 VKNNGTA -> NARGGRVR (IN REF. 3).
SQ SEQUENCE 416 AA; 44773 MW; AED793F7A560CB7 CRC64;
```

Query Match 74.4%; Score 29; DB 1; Length 416;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSLF 9
|:|:|:|:|
DB 290 GVNASSSRF 298

Search completed: January 29, 2004, 14:53:07
Job time : 6.8253 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: January 29, 2004, 14:46:19 : Search time 26.8676 Seconds
(without alignments)
86.441 Million cell updates/sec

Title: US-10-032-950-1
Perfect score: 39
Sequence: 1 GYVAXSLF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 segs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.23.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_orcanelle.*
9: sp_plant.*
10: sp_rodent.*
11: sp_virus.*
12: sp_virus.*
13: sp_virus.*
14: sp_virus.*
15: sp_virus.*
16: sp_virus.*
17: sp_virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	92.3	47	16	O33586
2	89.7	189	11	O32279
3	89.7	406	11	O8VH34
4	89.7	407	11	O8DC13
5	87.2	333	16	O34933
6	82.1	46	2	O9F1U5
7	82.1	895	10	O9SC24
8	82.1	895	10	O94C93
9	79.5	296	2	O47402
10	79.5	313	10	O8RT04
11	79.5	394	16	O8PM24
12	79.5	406	16	O9X5N8
13	79.5	408	16	O8EV24
14	79.5	412	16	O8FDP8
15	79.5	783	2	O9LAW0
16	76.9	109	17	O9UXG2

17	30	76.9	255	13	O985E3	O985E3 carassius a
18	30	76.9	360	16	O983P9	O983P9 rhizobium l
19	30	76.9	420	16	O8ZHN4	O8ZHN4 yerania pe
20	30	76.9	465	16	O8XHH9	O8XHH9 clostridium
21	30	76.9	465	5	O9VEI3	O9VEI3 drosophila
22	30	76.9	465	5	O960E7	O960E7 drosophila
23	30	76.9	560	10	O64883	O64883 arabidopsis
24	30	76.9	633	5	O18227	O18227 caenorhabdi
25	30	76.9	911	16	O8DXK6	O8DXK6 streptococc
26	29	74.4	63	2	O8RCX4	O8RCX4 streptococc
27	29	74.4	91	10	O8M459	O8M459 arabidopsis
28	29	74.4	119	10	O8H608	O8H608 oryza sativ
29	29	74.4	183	9	O8W760	O8W760 bacterioph
30	29	74.4	193	15	O9DV13	O9DV13 human immun
31	29	74.4	251	4	O96140	O96140 homo sapien
32	29	74.4	259	8	O8HN50	O8HN50 brugia mala
33	29	74.4	271	16	O8DPY3	O8DPY3 streptococc
34	29	74.4	311	4	O8BRD2	O8BRD2 homo sapien
35	29	74.4	335	4	O9NFI3	O9NFI3 homo sapien
36	29	74.4	366	2	O9KH93	O9KH93 escherichia
37	29	74.4	366	2	O9KH88	O9KH88 escherichia
38	29	74.4	366	2	O9XU22	O9XU22 escherichia
39	29	74.4	366	2	O9KH89	O9KH89 escherichia
40	29	74.4	366	2	O9KH92	O9KH92 escherichia
41	29	74.4	366	2	O9K2M3	O9K2M3 escherichia
42	29	74.4	366	2	O9K2H8	O9K2H8 escherichia
43	29	74.4	366	2	O9KH91	O9KH91 escherichia
44	29	74.4	378	2	O8KQ42	O8KQ42 pseudomonas
45	29	74.4	382	10	O8L898	O8L898 dunaliella

ALIGNMENTS

RESULT 1
ID O33586 PRELIMINARY: PRT: 47 AA.
AC O33586;
DT 01-JUN-1998 (TREMBLrel. 05, Created)
DT 01-JUN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE AGRD (AGRD protein).
GN AGRD OR SAV2037 OR SA1842.1 OR SAS066.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S. aureus; STRAIN=SA502A;
RX MEDLINE=97342847; PubMed=9197262;
RA U. G., Beavis R., Novick R.P.,
RT "Bacterial interference caused by autoinducing peptide variants.",
RL Science 276:2027-2030 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S. aureus (strain Mu50), and S. aureus (strain N315);
RX MEDLINE=2111952; PubMed=1148146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K., I., Nagai Y., Itoh T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani T., Takahashi N.K., Sawano T., Inoue R., I., Kato C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiratsuka K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus",
Lancet 357:1225-1240 (2001).
DR EMBL, AF001782; AAB63265.1;
DR EMBL, AF003364; BAB58199.1;
DR EMBL, AF003365; BAB58199.1;
KW Complete proteome.

SQ SEQUENCE 47 AA; 5149 MW; 22893764DF54BFA7 CRC64;

Query Match 92.3%; Score 36; DB 16; Length 47;
Best Local Similarity 88.9%; Pred. No. 0.67;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSLF 9
DB 24 GVNAXSSLF 32

RESULT 2

ID Q922T9 PRELIMINARY; PRT; 189 AA.

AC Q922T9; 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
OS Similar to lysosomal membrane glycoprotein 1 (fragment).
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_Taxid=10090;
RP SEQUENCE FROM N.A.
RA Strubeberg R.;
RT Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006785; AA06785.1;
DR InterPro; IPR002000; Lamp.
DR Pfam; PF01299; Lamp; 1.
DR PRINTS; PR00336; LYSASOCTDMP.
DR PROSITE; PS00310; LAMP_1; 1.
DR PROSITE; PS00311; LAMP_2; 1.
FT NON_TER 1
SQ SEQUENCE 189 AA; 20456 MW; 71F16D69BA4066FA CRC64;

Query Match 89.7%; Score 35; DB 11; Length 189;
Best Local Similarity 77.8%; Pred. No. 4.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSLF 9
DB 63 GVNAXSSLF 71

RESULT 3

ID Q8VH34 PRELIMINARY; PRT; 406 AA.

AC Q8VH34; 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_Taxid=10090;
RP SEQUENCE FROM N.A.
RA Ono K.; Han J.;
RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY069668; AAL58070.1;
DR InterPro; IPR002000; Lamp.
DR Pfam; PF01299; Lamp; 1.
DR PRINTS; PR00336; LYSASOCTDMP.
DR PROSITE; PS00310; LAMP_1; 2.
DR PROSITE; PS00311; LAMP_2; 1.
SQ SEQUENCE 406 AA; 43879 MW; C1BD373548ADFA85 CRC64;

Query Match 89.7%; Score 35; DB 11; Length 406;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSLF 9
DB 280 GVNAXSSLF 288

RESULT 4

ID Q9DC13 PRELIMINARY; PRT; 407 AA.

AC Q9DC13; 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
OS Lysosomal membrane glycoprotein 1.
GN LAMP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_Taxid=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085650; PubMed=11217851;
RA Kawai J.; Shinagawa A.; Shibata K.; Yoshino M.; Itoh M.; Ishii Y.;
RA Aizawa K.; Iizawa M.; Nishi K.; Kiyosawa H.; Kondo S.; Yamana K.;
RA Saito T.; Okazaki Y.; Gotohori T.; Boto H.; Kasukawa T.; Saito R.;
RA Kadota K.; Matsuda H.A.; Ashburner M.; Batalov S.; Casavant T.;
RA Flaischmann W.; Gaasterland T.; Glasl C.; King B.; Kochiwa H.;
RA Kuehl P.; Lewis S.; Matsuo Y.; Nikaido I.; Resole G.; Quackenbush J.;
RA Schmitt L.M.; Staubli F.; Suzuki R.; Tomita M.; Wagner L.; Washio T.;
RA Sakai K.; Okido T.; Furuno M.; Aono H.; Baldarelli R.; Barsh G.;
RA Blake J.; Boffelli D.; Bojunga N.; Carrinci P.; de Bonaldo M.F.;
RA Bromstein M.J.; Bulc C.; Fletcher C.; Fujita M.; Gariboldi M.;
RA Guerinich S.; Hill D.; Holman W.; Hume D.A.; Kamiya M.; Lee N.H.;
RA Lyons P.; Marchionni L.; Mashima Y.; Mazzarelli U.; Mombaerts P.;
RA Nordone P.; Ring B.; Ringwald M.; Rodriguez I.; Sakamoto N.;
RA Sasaki H.; Sato K.; Schoenbach C.; Seya T.; Shibata Y.; Storch K.-F.;
RA Suzuki H.; Toyooka K.; Wang K.H.; Weitz C.; Whitaker C.; Wilming L.;
RA Wyszewski-Boris A.; Yoshida K.; Hasegawa Y.; Kawai H.; Kohsaki S.;
RA Hayashizaki Y.;
RT "functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
DR EMBL; AK004637; BAB23428.1;
DR MGD; MGI:96745; Lamp1.
DR InterPro; IPR002000; Lamp.
DR Pfam; PF01299; Lamp; 1.
DR PRINTS; PR00336; LYSASOCTDMP.
DR PROSITE; PS00310; LAMP_1; 2.
DR PROSITE; PS00311; LAMP_2; 1.
SQ SEQUENCE 407 AA; 43936 MW; 696D0C79F627DA84 CRC64;

Query Match 89.7%; Score 35; DB 11; Length 407;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSLF 9
DB 281 GVNAXSSLF 289

RESULT 5

ID Q34933 PRELIMINARY; PRT; 333 AA.

AC Q34933; 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
GN YFMD protein.
GN YFMD.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OK NCBI_Taxid=1423;
RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Mosser I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertolo M.G., Bessières P., Bolotin A., Borcher S.,
 RA Borris R., Boursier L., Brans A., Brun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani U.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Dancourt F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Ertan K.D., Errington J., Fabre C., Ferrari E., Foulger D.,
 RA Filtz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,
 RA Gilm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Hachez J., Harwood C.R., Hentz A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasanara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Nockack M.,
 RA Moore D., O'Reilly M., Ogawa K., Ogikawa A., Oudega B., Park S.H.,
 RA Paro V., Pohl T.M., Portetelle D., Potwilk S., Prescott A.M.,
 RA Presean E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serior P., Shin B.S., Seldo B.,
 RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpsira P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandendol M., Vannier F., Vassart A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weltenegeger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.P., Zimstein E., Yoshikawa H., Dancin A.;
 RT "The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*";
 RT Nature 390:249-256(1997).
 RL Nature 390:249-256(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Kunst F., Ogasawara N., Yoshikawa H., Dancin A.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RC327;
 RX MEDLINE=97417488; PubMed=9272861;
 RA Yamamoto H., Uchiyama S., Nigroch F.A., Sekiguchi J.;
 RT "Cloning and sequencing of a 35.7 kb in the 70 degree-73 degree region
 of the *Bacillus subtilis* genome reveal genes for a new two-component
 RT system, three spore germination proteins, an iron uptake system and a
 RT general stress response protein.";
 RT Gene 194:191-199(1997).
 RL Gene 194:191-199(1997).
 DR EMBL; Z99108; CAB12580.1; -;
 DR EMBL; D86417; BAA23318.1; -;
 DR InterPro; IPR000522; FeccD.
 DR Pfam; PF01032; FeccD; 1.
 DR ProDom; PD001557; FeccD; 1.
 KW Complete proteome.
 SQ SEQUENCE 333 AA; 35098 MW; 9F4BE91B8DEB761 CRC64;
 QY Query Match 87.2%; Score 34; DB 16; Length 333;
 Best Local Similarity 77.8%; Pred. No. 15;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Db 99 GVNAGASLF 107

DE AGRD.
 GN AGRD.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxId=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=125;
 RX MEDLINE=21137858; PubMed=11240104;
 RA Takeuchi S., Maeda T., Hashimoto N., Imaizumi K., Kaidoh T.,
 RA Hayakawa Y.;
 RT "Variation of the *agr* locus in *Staphylococcus aureus* isolates from
 RT cows with mastitis.";
 RL Vet. Microbiol. 79:267-274(2001).
 DR EMBL; AB043554; BAB18547.1; -;
 SQ SEQUENCE 46 AA; 5050 MW; 404714CD48FA77E CRC64;
 QY Query Match 82.1%; Score 32; DB 2; Length 46;
 Best Local Similarity 77.8%; Pred. No. 5.5;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Db 24 GVNACGIF 32

RESULT 7
 Q9SCZ4 PRELIMINARY; PRT; 895 AA.
 ID Q9SCZ4;
 AC Q9SCZ4;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Receptor-protein kinase-like protein.
 GN F26013.190.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxId=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Delany M., Berger C., Cooke R., Grellier F., Laudie M., Mewes H.W.,
 RA Lemcke K., Mayer K.F.X., Queller M., Salanoubat M.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 CC -1; SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; A113452; CAB63019.1; -;
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_Thr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP_1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM_1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST_1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 895 AA; 98149 MW; CFAEFB8E314BCC6 CRC64;
 QY Query Match 82.1%; Score 32; DB 10; Length 895;
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Db 122 GVNATNSLF 130

RESULT 8
 Q94C93 PRELIMINARY; PRT; 895 AA.
 ID Q94C93

AC Q94C93;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical 98.2 kDa protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosidia II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
 RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
 RA Jones T., Kamuya A., Karlin-Neumann G., Kawai J., Kim C., Koeseema E.,
 RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
 RA Palm C.J., Sakurai T., Satou M., Seki M., Shin P., Southwick A.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,
 RT "Full length cDNA of gene F26013.190/AT1G51550 (GI:6572076)."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1 - SIMILARITY: BELONGS TO THE SSR/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AY035053; AKS9558.1; -
 DR InterPro; IPR000719; Prot Kinase.
 DR InterPro; IPR002290; Ser Thr_pkinase.
 DR Pfam; PF00069; pkinase; 1
 DR PRODOM; PD000001; Prot Kinase; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS0108; PROTEIN_KINASE_ST; 1.
 KW Hypothetical protein; ATP-binding; Kinase;
 KW Serine/threonine-protein kinase; Transferrase.
 SQ SEQUENCE 895 AA; 98191 MW; 200DB3B35971637 CRC64;

Query Match
 Best Local Similarity 82.1%; Score 32; DB 10; Length 895;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GVNAXSSLF 9
 DB 122 GVNAXSSLF 130

RESULT 9
 Q47402 PRELIMINARY; PRT; 295 AA.

AC Q47402;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE KpsS protein (Fragment).
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=K-12/K-1;
 RX MEDLINE=92138601; PubMed=1735705;
 RT "Functional analysis of the stalyltransferase complexes in Escherichia coli K1 and K92."
 RT J. Bacteriol. 174:1099-1108(1992).
 DR EMBL; M76370; AAA24214.1; -
 FT NON-TER
 SQ SEQUENCE 296 AA; 35733 MW; AEEBA3DA46AC308B CRC64;

Query Match
 Best Local Similarity 79.5%; Score 31; DB 2; Length 296;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GVNAXSSL 8

DB 22 GVNAXSSL 29

RESULT 10

Q8RU04 PRELIMINARY; PRT; 313 AA.

AC Q8RU04;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE OSUNBA0026J14.9 protein (OSUNBA0051H17.31 protein).
 GN OSUNBA0026J14.9 OR OSUNBA0051H17.31.
 OS Oryza sativa (Japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrharioideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa (Japonica cultivar-group) genomic DNA, chromosome 1, BAC clone:OSUNBA0026J14.";
 RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa (Japonica cultivar-group) genomic DNA, chromosome 1, BAC clone:OSUNBA0051H17.";
 RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP004231; BAB99223.1; -
 DR EMBL; AP004232; BAB90716.1; -
 DR Gramene; Q8RU04; -
 SQ SEQUENCE 313 AA; 35389 MW; 94C14BB876201E42 CRC64;

Query Match
 Best Local Similarity 79.5%; Score 31; DB 10; Length 313;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVNAXSSLF 9
 DB 102 GVNAXSSLF 110

RESULT 11

Q9PMP4 PRELIMINARY; PRT; 394 AA.

AC Q9PMP4;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Possible polysaccharide modification protein.
 GN Cj1413C.

OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Campylobacteraceae; Campylobacter.
 OX NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCCTC 11168;
 RX MEDLINE=20150912; PubMed=10668204;
 RA Parkhill J., Wren B.W., Mungall K., Kellley J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
 RA Jørgensen K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Randleman M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Bartell B.G.;
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences.";
 RT Nature 403:665-668(2000).
 DR EMBL; AL139078; CAB73837.1; -
 KM Complete proteome.

SQ SEQUENCE 394 AA; 46938 MW; E4293DB5879FD64E CRC64;
 Query Match 79.5%; Score 31; DB 16; Length 394;
 Best Local Similarity 87.5%; Pred. No. 88;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSL 8
 DB 126 GVNAXSSL 133

RESULT 12
 Q9X5N8 PRELIMINARY; PRT; 406 AA.
 AC Q9X5N8;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Capsule export protein KpsS.
 GN KPS.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D114;
 RX MEDLINE=9194741; PubMed=10094710;
 RA Clarke B.R., Pearce R., Roberts I.S.;
 RT "Genetic organization of the Escherichia coli K10 capsule gene
 cluster: identification and characterization of two conserved regions
 in group III capsule gene clusters encoding polysaccharide transport
 functions."
 RT J. Bacteriol. 181:2279-2285(1999).
 RL EMBL; AF127177; AAD32184.1; -
 DR SEQUENCE 406 AA; 48449 MW; 99C851E798C3DC0B CRC64;

Query Match 79.5%; Score 31; DB 2; Length 406;
 Best Local Similarity 87.5%; Pred. No. 91;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSL 8
 DB 127 GVNAXSSL 134

RESULT 13
 Q8EV24 PRELIMINARY; PRT; 408 AA.
 AC Q8EV24;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Amino transferase NIFS.
 GN MYE7430.
 OS Mycoplasma penetrans.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=28227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HF-2;
 RX MEDLINE=2354719; PubMed=12466555;
 RA Sasaki Y., Ishikawa T., Yamashita A., Oshima K., Kanai T., Furuya K.,
 RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
 RT "The complete genomic sequence of Mycoplasma penetrans, an
 RT intracellular bacterial pathogen in humans."
 RT Nucleic Acids Res. 30:5293-5300(2002).
 RL EMBL; AF004173; BAC44537.1; -
 DR TRANSFERASE; Complete proteome.
 KW SEQUENCE 408 AA; 46412 MW; C406928B08DB648D CRC64;

Query Match 79.5%; Score 31; DB 16; Length 408;
 Best Local Similarity 55.6%; Pred. No. 91;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSLF 9
 DB 326 GVNAXSLF 334

RESULT 14
 Q8FDP8 PRELIMINARY; PRT; 412 AA.
 AC Q8FDP8;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Kps protein.
 GN KPS OR C3691.
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RX MEDLINE=2238234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G., III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 RT of uropathogenic Escherichia coli."
 RT Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 RL EMBL; AE016766; AAN82139.1; -
 DR Complete proteome.
 KW SEQUENCE 412 AA; 48808 MW; A9DE2A6F3E144C94 CRC64;

Query Match 79.5%; Score 31; DB 16; Length 412;
 Best Local Similarity 87.5%; Pred. No. 92;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSL 8
 DB 127 GVNAXSSL 134

RESULT 15
 Q9LAW0 PRELIMINARY; PRT; 783 AA.
 AC Q9LAW0;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Phase-variable hemagglutinin.
 GN VHA2.
 OS Mycoplasma synoviae.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2109;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WTU-1853;
 RX MEDLINE=20156683; PubMed=10692167;
 RA Noorichammadi A.H., Markham P.F., Kanci A., Whithear K.G.,
 RA Browning G.F.;
 RT "A novel mechanism for control of antigenic variation in the
 RT haemagglutinin gene family of Mycoplasma synoviae."
 RT Mol. Microbiol. 35:911-923(2000).
 RL EMBL; AF088597; AAD42254.1; -
 DR SEQUENCE 783 AA; 84326 MW; A1C86C5417B8690C CRC64;

Query Match 79.5%; Score 31; DB 2; Length 783;
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSLF 9

Fri Jan 30 09:06:38 2004

us-10-032-950-1.rpt

Page 6

Db 653 GHAESLIF 661

Search completed: January 29, 2004, 14:56:43
Job time : 27.8676 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:52:19 ; Search time 11.1176 Seconds
(without alignments)
34.252 Million cell updates/sec

Title: US-10-032-950-1

Perfect score: 39

Sequence: 1 GVNAXSSLF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/aa/6C.COMB.pep.*
6: /cgn2_6/ptodata/1/aa/6D.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	94.9	9	US-09-339-511-1	Sequence 1, Appl
2	36	92.3	9	US-08-861-476C-3	Sequence 3, Appl
3	36	92.3	47	US-08-861-476C-6	Sequence 6, Appl
4	34	87.2	9	US-09-339-511-5	Sequence 5, Appl
5	34	87.2	9	US-09-339-511-6	Sequence 6, Appl
6	33	84.6	9	US-09-339-511-2	Sequence 2, Appl
7	32	82.1	9	US-09-339-511-7	Sequence 7, Appl
8	31	79.5	9	US-09-339-511-4	Sequence 4, Appl
9	29	74.4	9	US-09-339-511-3	Sequence 3, Appl
10	29	74.4	332	US-09-882-835-2	Sequence 2, Appl
11	29	74.4	353	US-08-073-807A-17	Sequence 17, Appl
12	29	74.4	380	US-08-073-807A-18	Sequence 18, Appl
13	29	74.4	382	US-08-186-833-2	Sequence 2, Appl
14	29	74.4	411	US-08-399-561-2	Sequence 2, Appl
15	29	74.4	416	US-08-073-807A-2	Sequence 2, Appl
16	29	74.4	416	US-09-311-784A-26	Sequence 26, Appl
17	29	74.4	1054	US-09-516-467-4	Sequence 4, Appl
18	29	74.4	1054	US-07-934-374-4	Sequence 4, Appl
19	29	74.4	1054	US-07-783-861C-2	Sequence 2, Appl
20	28	71.8	69	US-09-134-001C-1105	Sequence 4105, Ap
21	28	71.8	465	US-09-252-991A-29387	Sequence 29387, A
22	27	69.2	264	US-09-053-197A-8	Sequence 8, Appl
23	27	69.2	264	US-09-085-761A-8	Sequence 8, Appl
24	27	69.2	293	US-09-252-991A-20791	Sequence 20791, A
25	27	69.2	319	US-09-724-510-2	Sequence 2, Appl
26	27	69.2	319	US-09-723-216-2	Sequence 2, Appl
27	27	69.2	319	US-09-675-227-2	Sequence 2, Appl

28	27	69.2	322	US-09-188-930-141	Sequence 141, App
29	27	69.2	322	US-09-312-283C-141	Sequence 141, App
30	27	69.2	339	US-09-107-832A-4904	Sequence 4904, Ap
31	27	69.2	343	US-09-967-908A-8	Sequence 8, Appl
32	27	69.2	348	US-09-252-991A-31001	Sequence 31001, A
33	27	69.2	363	US-09-967-908A-10	Sequence 10, Appl
34	27	69.2	373	US-09-967-908A-4	Sequence 4, Appl
35	27	69.2	381	US-09-198-452A-193	Sequence 193, App
36	27	69.2	416	US-09-188-930-341	Sequence 341, App
37	27	69.2	431	US-09-312-283C-341	Sequence 341, App
38	27	69.2	464	US-09-252-991A-27559	Sequence 27559, A
39	27	69.2	464	US-09-252-991A-28048	Sequence 28048, A
40	27	69.2	536	US-09-595-857B-31	Sequence 31, Appl
41	27	69.2	614	US-09-832-614A-2	Sequence 2, Appl
42	27	69.2	614	US-09-832-614A-2	Sequence 2, Appl
43	27	69.2	791	US-08-394-880B-2	Sequence 2, Appl
44	27	69.2	1024	US-09-091-117-5	Sequence 5, Appl
45	27	69.2	1024	US-09-091-117-5	Sequence 5, Appl

ALIGNMENTS

```

RESULT 1
; Sequence 1, Application US/09339511
; Patent No. 6337385
; GENERAL INFORMATION:
; APPLICANT: Mult, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. 6337385ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/09/339,511
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: 60/090,402
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-1
Query Match          94.9%; Score 37; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GVNAXSSLF 9
Db      1 GVNAXSSLF 9
RESULT 2
; Sequence 3, Application US/08861476C
; Patent No. 6447786
; GENERAL INFORMATION:
; APPLICANT: New York University Medical Center
; TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
; FILE REFERENCE: 63753/7
; CURRENT APPLICATION NUMBER: US/08/861,476C
; CURRENT FILING DATE: 1997-05-22

```

/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO: 3
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
US-08-861-476C-3

Query Match 92.3%; Score 36; DB 4; Length 9;
Best Local Similarity 88.9%; Pred. No. 2.5e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSLF 9
|||
Db 1 GVNACSSLF 9

RESULT 3
US-08-861-476C-6
/ Sequence 6, Application US/08861476C
/ Patent No. 6447786
/ GENERAL INFORMATION:
/ APPLICANT: New York University Medical Center
/ TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
/ FILE REFERENCE: 63753/7
/ CURRENT APPLICATION NUMBER: US/08/861,476C
/ CURRENT FILING DATE: 1997-05-22
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO: 6
/ LENGTH: 47
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
US-08-861-476C-6

Query Match 92.3%; Score 36; DB 4; Length 47;
Best Local Similarity 88.9%; Pred. No. 0.25;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSLF 9
|||
Db 24 GVNACSSLF 32

RESULT 4
US-09-339-511-5
/ Sequence 5, Application US/09339511
/ Patent No. 6337385
/ GENERAL INFORMATION:
/ APPLICANT: Muir, Tom
/ APPLICANT: Mayville, Patricia
/ APPLICANT: No. 6337385ick, Richard P.
/ APPLICANT: Beavis, Ronald
/ APPLICANT: Ji, Guangyong
/ TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
/ FILE REFERENCE: 600-1-231N
/ CURRENT APPLICATION NUMBER: US/09/339,511
/ CURRENT FILING DATE: 1999-06-24
/ PRIOR APPLICATION NUMBER: 60/090,402
/ PRIOR FILING DATE: 1998-06-24
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO: 5
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic
/ NAME/KEY: VARIANT
/ LOCATION: (5)
/ OTHER INFORMATION: Xaa represents any amino acid at this position.

US-09-339-511-5

Query Match 87.2%; Score 34; DB 4; Length 9;
Best Local Similarity 86.9%; Pred. No. 2.5e+05;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVNAXSLF 9
|||
Db 1 GVNAXSLF 9

RESULT 5
US-09-339-511-6
/ Sequence 6, Application US/09339511
/ Patent No. 6337385
/ GENERAL INFORMATION:
/ APPLICANT: Muir, Tom
/ APPLICANT: Mayville, Patricia
/ APPLICANT: No. 6337385ick, Richard P.
/ APPLICANT: Beavis, Ronald
/ APPLICANT: Ji, Guangyong
/ TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
/ FILE REFERENCE: 600-1-231N
/ CURRENT APPLICATION NUMBER: US/09/339,511
/ CURRENT FILING DATE: 1999-06-24
/ PRIOR APPLICATION NUMBER: 60/090,402
/ PRIOR FILING DATE: 1998-06-24
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO: 6
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic
/ NAME/KEY: VARIANT
/ LOCATION: (5)
/ OTHER INFORMATION: Xaa represents any amino acid at this position.

Query Match 87.2%; Score 34; DB 4; Length 9;
Best Local Similarity 86.9%; Pred. No. 2.5e+05;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVNAXSLF 9
|||
Db 1 GVNAXSLF 9

RESULT 6
US-09-339-511-2
/ Sequence 2, Application US/09339511
/ Patent No. 6337385
/ GENERAL INFORMATION:
/ APPLICANT: Muir, Tom
/ APPLICANT: Mayville, Patricia
/ APPLICANT: No. 6337385ick, Richard P.
/ APPLICANT: Beavis, Ronald
/ APPLICANT: Ji, Guangyong
/ TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
/ FILE REFERENCE: 600-1-231N
/ CURRENT APPLICATION NUMBER: US/09/339,511
/ CURRENT FILING DATE: 1999-06-24
/ PRIOR APPLICATION NUMBER: 60/090,402
/ PRIOR FILING DATE: 1998-06-24
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO: 2
/ LENGTH: 9
/ TYPE: PRT

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
; US-09-339-511-2

Query Match
Best Local Similarity 84.6%; Score 33; DB 4; Length 9;
Pred. No. 2.5e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYNAXSRLF 9
Db 1 GYNAXSRLF 9

RESULT 7
; US-09-339-511-7
; Sequence 7, Application US/09339511
; Patent No. 6337385
; GENERAL INFORMATION:
; APPLICANT: Muir, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. 6337385ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/09/339,511
; CURRENT FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
; US-09-339-511-7

Query Match
Best Local Similarity 82.1%; Score 32; DB 4; Length 9;
Pred. No. 2.5e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYNAXSRLF 9
Db 1 GYNAXSRLF 9

RESULT 8
; US-09-339-511-4
; Sequence 4, Application US/09339511
; Patent No. 6337385
; GENERAL INFORMATION:
; APPLICANT: Muir, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. 6337385ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/09/339,511
; CURRENT FILING DATE: 1999-06-24
```

```

; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
; US-09-339-511-4

Query Match
Best Local Similarity 79.5%; Score 31; DB 4; Length 9;
Pred. No. 2.5e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VNAXSRLF 9
Db 2 VNAXSRLF 9

RESULT 9
; US-09-339-511-3
; Sequence 3, Application US/09339511
; Patent No. 6337385
; GENERAL INFORMATION:
; APPLICANT: Muir, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. 6337385ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/09/339,511
; CURRENT FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
; US-09-339-511-3

Query Match
Best Local Similarity 74.4%; Score 29; DB 4; Length 9;
Pred. No. 2.5e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYNAXSRLF 9
Db 1 GYNAXSRLF 9

RESULT 10
; US-09-882-835-2
; Sequence 2, Application US/09882835
; Patent No. 6462187
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: A NOVEL HUMAN THIOREDOXIN FAMILY
; MEMBER AND USES THEREOF
```

FILE REFERENCE: 10448-063001
CURRENT APPLICATION NUMBER: US/09/882,835
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/211,673
PRIOR FILING DATE: 2000-06-15
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 2
LENGTH: 332
TYPE: PRT
ORGANISM: Homo sapiens
US-09-882-835-2

Query Match 74.4% Score 29; DB 4; Length 332;
Best Local Similarity 66.7% Pred. No. 87;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVNAXSLF 9
Db 197 GVNAXSLF 205

RESULT 11

US-08-073-807A-17
Sequence 17, Application US/08073807A
Patent No. 5646248
GENERAL INFORMATION:
APPLICANT: Sawada, Ritsuko
APPLICANT: Lowe, John B.
APPLICANT: Fukuda, Minoru
TITLE OF INVENTION: CELL SURFACE LAMP EXPRESSION AND
TITLE OF INVENTION: SELECTIN-DEPENDENT ADHESION
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/073,807A
FILING DATE: 08-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9567
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-8949
TELEFAX: (619) 535-9001
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 353 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-073-807A-17

Query Match 74.4% Score 29; DB 1; Length 353;
Best Local Similarity 66.7% Pred. No. 93;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVNAXSLF 9
Db 263 GVNAXSLF 271

RESULT 12

US-08-073-807A-18

Sequence 18, Application US/08073807A
Patent No. 5646248
GENERAL INFORMATION:
APPLICANT: Sawada, Ritsuko
APPLICANT: Lowe, John B.
APPLICANT: Fukuda, Minoru
TITLE OF INVENTION: CELL SURFACE LAMP EXPRESSION AND
TITLE OF INVENTION: SELECTIN-DEPENDENT ADHESION
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/073,807A
FILING DATE: 08-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9567
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-073-807A-18

Query Match 74.4% Score 29; DB 1; Length 380;
Best Local Similarity 66.7% Pred. No. 18+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVNAXSLF 9
Db 290 GVNAXSLF 298

RESULT 13

US-08-186-833-2
Sequence 2, Application US/08186833
Patent No. 5563324
GENERAL INFORMATION:
APPLICANT: Tarczyński, Mitchell C.
APPLICANT: Jensen, Richard G.
APPLICANT: Bonhert, Hans J.
TITLE OF INVENTION: Transgenic Plants with Enhanced Mannitol
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles & Brady
STREET: P.O. Box 2113, First Wisconsin Plaza
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701-2113

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word, Version #5.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/186,833

FILING DATE:
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/871,416
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Seay, Nicholas J.
 REGISTRATION NUMBER: 27386
 REFERENCE/DOCKET NUMBER: 9221490026
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (608)251-5000
 TELEFAX: (608) 251-9166
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 382 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-186-833-2

Query Match 74.4%; Score 29; DB 1; Length 382;
 Best Local Similarity 75.0%; Pred. No. 1e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

CY 1 GVNAXSSL 8
 DB 62 GVNAXSSI 69

RESULT 14
 US-08-399-561-2
 Sequence 2, Application US/08399561
 Patent No. 5792903
 GENERAL INFORMATION:
 APPLICANT: Hirschberg, Joseph
 APPLICANT: Cunningham Jr., Francis X.
 APPLICANT: Gantt, Elisabeth
 TITLE OF INVENTION: Lycogene Cyclase Gene
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Kohn & Associates
 STREET: 30500 No. 5792903Western Highway, Suite 410
 CITY: Farmington Hills
 STATE: MI
 COUNTRY: US
 ZIP: 48334
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/399,561
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Kohn, Kenneth I.
 REGISTRATION NUMBER: 30,955
 REFERENCE/DOCKET NUMBER: P-305 (Hebrew Univ.)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 810-539-5050
 TELEFAX: 810-539-5055
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 411 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-399-561-2

Query Match 74.4%; Score 29; DB 1; Length 411;
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

CY 1 GVNAXSSL 8
 DB 299 GVNAXSSL 306

RESULT 15
 US-08-073-807A-2
 Sequence 2, Application US/08073807A
 Patent No. 5646248
 GENERAL INFORMATION:
 APPLICANT: Sawada, Ritsuko
 APPLICANT: Lowe, John B.
 APPLICANT: Fukuda, Minoru
 TITLE OF INVENTION: CELL SURFACE LAMP EXPRESSION AND
 TITLE OF INVENTION: SELECTIN-DEPENDENT ADHESION
 NUMBER OF SEQUENCES: 18
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Campbell and Flores
 STREET: 4370 La Jolla Village Drive, Suite 700
 CITY: San Diego
 STATE: California
 COUNTRY: United States
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/073,807A
 FILING DATE: 08-JUN-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-LJ 9567
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 416 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-073-807A-2

Query Match 74.4%; Score 29; DB 1; Length 416;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 CY 1 GVNAXSLF 9
 DB 290 GVNAXSSRF 298

Search completed: January 29, 2004, 15:03:20
 Job time : 12.1176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:47:39 ; Search time 25.5441 Seconds
(without alignments)
73.223 Million cell updates/sec

Title: US-10-032-950-1

Sequence: 1 GYNAXSLF 9

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched:

789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters:

789580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA:*

1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/prodata/1/pubpaa/US07_PUB_PUB.pep:*

3: /cgn2_6/prodata/1/pubpaa/US06_PUB_PUB.pep:*

4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/prodata/1/pubpaa/US07_PUB_PUB.pep:*

6: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*

7: /cgn2_6/prodata/1/pubpaa/US08_PUB_PUB.pep:*

8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*

10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*

11: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*

12: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*

13: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*

14: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*

15: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*

16: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*

17: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep:*

18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	94.9	9	US-10-032-950-1	Sequence 1, App1
2	36	92.3	9	US-10-201-444-3	Sequence 3, App1
3	36	92.3	47	US-10-201-444-5	Sequence 6, App1
4	34	87.2	9	US-10-032-950-5	Sequence 5, App1
5	34	87.2	9	US-10-032-950-6	Sequence 6, App1
6	33	84.6	9	US-10-032-950-2	Sequence 7, App1
7	32	82.1	9	US-10-032-950-7	Sequence 7, App1
8	31	79.5	9	US-10-032-950-4	Sequence 4, App1
9	30	76.9	455	US-10-004-378A-130	Sequence 130, App
10	30	76.9	455	US-10-004-378A-131	Sequence 131, App
11	30	76.9	455	US-10-108-605-227	Sequence 227, App
12	30	76.9	633	US-10-369-493-6225	Sequence 6225, App
13	29	74.4	9	US-10-032-950-3	Sequence 3, App1
14	29	74.4	332	US-10-145-586-46	Sequence 46, App1
15	29	74.4	382	US-09-815-242-10385	Sequence 10385, A

16	29	74.4	416	US-10-371-069-26	Sequence 26, App1
17	29	74.4	416	US-10-371-645-26	Sequence 26, App1
18	29	74.4	416	US-10-371-645-26	Sequence 26, App1
19	29	74.4	490	US-10-243-601-3	Sequence 3, App1
20	29	74.4	490	US-10-243-601-3	Sequence 3, App1
21	29	74.4	605	US-10-100-303A-110	Sequence 110, App
22	29	74.4	747	US-09-978-855A-459	Sequence 459, App
23	29	74.4	747	US-09-978-855A-459	Sequence 459, App
24	29	74.4	747	US-09-978-855A-459	Sequence 459, App
25	29	74.4	747	US-09-978-855A-459	Sequence 459, App
26	29	74.4	747	US-09-978-855A-459	Sequence 459, App
27	29	74.4	747	US-09-978-855A-459	Sequence 459, App
28	29	74.4	747	US-09-978-855A-459	Sequence 459, App
29	29	74.4	747	US-09-978-855A-459	Sequence 459, App
30	29	74.4	747	US-09-978-855A-459	Sequence 459, App
31	29	74.4	747	US-09-978-855A-459	Sequence 459, App
32	29	74.4	747	US-09-978-855A-459	Sequence 459, App
33	29	74.4	747	US-09-978-855A-459	Sequence 459, App
34	29	74.4	747	US-09-978-855A-459	Sequence 459, App
35	29	74.4	747	US-09-978-855A-459	Sequence 459, App
36	29	74.4	747	US-09-978-855A-459	Sequence 459, App
37	29	74.4	747	US-09-978-855A-459	Sequence 459, App
38	29	74.4	747	US-09-978-855A-459	Sequence 459, App
39	29	74.4	747	US-09-978-855A-459	Sequence 459, App
40	29	74.4	747	US-09-978-855A-459	Sequence 459, App
41	29	74.4	747	US-09-978-855A-459	Sequence 459, App
42	29	74.4	747	US-09-978-855A-459	Sequence 459, App
43	29	74.4	747	US-09-978-855A-459	Sequence 459, App
44	29	74.4	747	US-09-978-855A-459	Sequence 459, App
45	29	74.4	747	US-09-978-855A-459	Sequence 459, App

ALIGNMENTS

RESULT 1

US-10-032-950-1

Sequence 1, Application US/10032950

Publication No. US20020077453A1

GENERAL INFORMATION:

APPLICANT: Mult, Tom

APPLICANT: Mayville, Patricia

APPLICANT: No. US20020077453A1, Richard P.

APPLICANT: Beavis, Ronald

APPLICANT: Ji, Guangyong

TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL

TITLE OF INVENTION: INTERFERENCE

FILE REFERENCE: 600-1-231N

CURRENT APPLICATION NUMBER: US/10/032,950

CURRENT FILING DATE: 2001-12-27

PRIOR APPLICATION NUMBER: 60/090,402

PRIOR FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 8

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 1

LENGTH: 9

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

OTHER INFORMATION: peptide

NAME/KEY: VARIANT

LOCATION: (5)

OTHER INFORMATION: Xaa represents any amino acid at this position.

US-10-032-950-1

Query Match 94.9% Score 37; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 76-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

QY 1 GYNAXSLF 9
1 GYNAXSLF 9

Db 1 GYNAXSLF 9

RESULT 2
US-10-201-444-3
; Sequence 3, Application US/10201444
; Publication No. US20030078378A1
; GENERAL INFORMATION:
; APPLICANT: New York University Medical Center
; TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
; FILE REFERENCE: 63753/7
; CURRENT APPLICATION NUMBER: US/10/201,444
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/08/861,476
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-201-444-3

Query Match 92.3%; Score 36; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 7e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSLSLF 9
Db 1 GVNACSSLF 9

RESULT 3
US-10-201-444-6
; Sequence 6, Application US/10201444
; Publication No. US20030078378A1
; GENERAL INFORMATION:
; APPLICANT: New York University Medical Center
; TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
; FILE REFERENCE: 63753/7
; CURRENT APPLICATION NUMBER: US/10/201,444
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/08/861,476
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-201-444-6

Query Match 92.3%; Score 36; DB 15; Length 47;
Best Local Similarity 88.9%; Pred. No. 0.44;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSLSLF 9
Db 24 GVNACSSLF 32

RESULT 4
US-10-032-950-5
; Sequence 5, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mult, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N

; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-5

Query Match 87.2%; Score 34; DB 13; Length 9;
Best Local Similarity 88.9%; Pred. No. 7e+05;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVNAXSLSLF 9
Db 1 GVNAXSLSLF 9

RESULT 5
US-10-032-950-6
; Sequence 6, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mult, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position..
US-10-032-950-6

Query Match 87.2%; Score 34; DB 13; Length 9;
Best Local Similarity 88.9%; Pred. No. 7e+05;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVNAXSLSLF 9
Db 1 GVNAXSLSLF 9

RESULT 6
US-10-032-950-2
; Sequence 2, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mult, Tom

```

; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-2

```

```

Query Match      84.6%; Score 33; DB 13; Length 9;
Best Local Similarity 88.9%; Pred. No. 7e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 GVNXSSLF 9
   |||||
Db 1 GVNXSSLF 9

```

```

RESULT 7
US-10-032-950-7
; Sequence 7, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Muir, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-7

```

```

Query Match      82.1%; Score 32; DB 13; Length 9;
Best Local Similarity 88.9%; Pred. No. 7e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 GVNXSSLF 9
   |||||
Db 1 GVNXSSAF 9

```

```

RESULT 8
US-10-032-950-4
; Sequence 4, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Muir, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-4

```

```

Query Match      79.5%; Score 31; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 VNXSSLF 9
   |||||
Db 2 VNXSSLF 9

```

```

RESULT 9
US-10-004-376A-130
; Sequence 130, Application US/10004376A
; Publication No. US20030228301A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Furtak, Kazaryna
; APPLICANT: Perna, Amanda
; APPLICANT: Patrusajan, Meera
; APPLICANT: Shimkels, Richard A
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Casman, Stacie J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Vernet, Corrine A
; APPLICANT: Spyrek, Kimberly A
; APPLICANT: Agee, Michele
; APPLICANT: Rastelli, Luca
; APPLICANT: Skeney, Suresh G
; APPLICANT: Grose, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Gerlach, Valerie
; APPLICANT: Bainger, Schomic
; APPLICANT: Macdougall, John R
; APPLICANT: Peyman, John A
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David J
; APPLICANT: Ellerman, Karen
; APPLICANT: Gangolli, Esha A
; TITLE OF INVENTION: No. US20030228301A1 Human Proteins, Polynucleotides Encoding

```

```

; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 21402-179
; CURRENT APPLICATION NUMBER: US/10/004,378A
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/242,882
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,765
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/300,206
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/242,789
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,768
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,767
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/243,622
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/273,047
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/243,591
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,950
; PRIOR FILING DATE: 2000-10-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 130
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-004-378A-130
```

```

Query Match          76.9%; Score 30; DB 12; Length 465;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 GVNXXSLF 9
Db 339 GVNLFSSIF 347
```

```

RESULT 10
US-10-004-378A-131
; Sequence 131, Application US/10004378A
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Furtak, Kazaryna
; APPLICANT: Perna, Amanda
; APPLICANT: Patuturajan, Neera
; APPLICANT: Shinkels, Richard A
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Casman, Stacie J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Malyanekar, Uriel M
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Verneq, Corrine A
; APPLICANT: Spytek, Kimberli A
; APPLICANT: Agee, Michele
; APPLICANT: Rastelli, Luca
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Gerlach, Valerie
; APPLICANT: Edinger, Schlomit
; APPLICANT: MacDougall, John R
; APPLICANT: Peyman, John A
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David J
; APPLICANT: Ellerman, Karen
; APPLICANT: Gangoli, Esha A
```

```

; TITLE OF INVENTION: No. US20030228301A1el Human Proteins, Polynucleotides Encoding
; FILE REFERENCE: 21402-179
; CURRENT APPLICATION NUMBER: US/10/004,378A
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/242,882
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,765
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/300,206
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/242,789
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,768
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,767
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/243,622
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/273,047
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/243,591
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,950
; PRIOR FILING DATE: 2000-10-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 131
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-004-378A-131
```

```

Query Match          76.9%; Score 30; DB 12; Length 465;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 GVNXXSLF 9
Db 339 GVNLFSSIF 347
```

```

RESULT 11
US-10-108-605-227
; Sequence 227, Application US/10108605
; Publication No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Steam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCOE
; FILE REFERENCE: 31133B
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 227
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-227
```

```

Query Match          76.9%; Score 30; DB 14; Length 465;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

QY 1 GYMAXSLF 9
 DB 339 GYVLFSSIF 347

RESULT 12

US-10-369-493-6225
 ; Sequence 6225, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 4/374
 ; SEQ ID NO 6225
 ; LENGTH: 633
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-10-369-493-6225

Query Match 76.9%; Score 30; DB 12; Length 633;
 Best Local Similarity 75.0%; Pred. No. 2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VMAXSLF 9
 DB 328 INAFSLF 335

RESULT 13
 US-10-032-950-3
 ; Sequence 3, Application US/10032950
 ; Publication No. US20020077453A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mulf, Tom
 ; APPLICANT: Mayville, Patricia
 ; APPLICANT: No. US20020077453A1ick, Richard P.
 ; APPLICANT: Beavis, Ronald
 ; APPLICANT: Ji, Guangyong
 ; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
 ; TITLE OF INVENTION: INTERFERENCE
 ; FILE REFERENCE: 600-1-231N
 ; CURRENT APPLICATION NUMBER: US/10/032,950
 ; CURRENT FILING DATE: 2001-12-27
 ; PRIOR APPLICATION NUMBER: 60/090,402
 ; PRIOR FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; OTHER INFORMATION: peptide
 ; NAME/KEY: VARIANT
 ; LOCATION: (5)
 ; OTHER INFORMATION: Xaa represents any amino acid at this position.
 US-10-032-950-3

Query Match 74.4%; Score 29; DB 13; Length 9;
 Best Local Similarity 88.9%; Pred. No. 7e+05;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYMAXSLF 9
 DB 1 GYMAXSLF 9

RESULT 14
 US-10-145-586-46
 ; Sequence 46, Application US/10145586
 ; Publication No. US20030138890A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandra Glucksmann, Maria
 ; APPLICANT: Silos-Santiago, Inmaculada
 ; APPLICANT: M. Galvin, Katherine
 ; APPLICANT: Welch, Nadine
 ; APPLICANT: Curtis, Rory A.J.
 ; APPLICANT: Bandaru, Rajasekhar
 ; APPLICANT: Kapeller-Liebermann, Rosana
 ; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTOR FAMILY MEMBERS,
 ; TITLE OF INVENTION: HUMAN THIOREDOXIN FAMILY MEMBERS, HUMAN LECITHINE-RICH
 ; TITLE OF INVENTION: REPEAT FAMILY MEMBERS, AND HUMAN RINGFINGER FAMILY MEMBER.
 ; FILE REFERENCE: 10448-188001
 ; CURRENT APPLICATION NUMBER: US/10/145,586
 ; CURRENT FILING DATE: 2002-05-14
 ; PRIOR APPLICATION removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 95
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 46
 ; LENGTH: 332
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-145-586-46

Query Match 74.4%; Score 29; DB 12; Length 332;
 Best Local Similarity 66.7%; Pred. No. 1.7e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYMAXSLF 9
 DB 197 GYVSLF 205

RESULT 15
 US-09-815-242-10385
 ; Sequence 10385, Application US/09815242
 ; Patent No. US20020061589A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Olsen, Karl L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Treweek, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; TITLE OF INVENTION: Prokaryotes
 ; FILE REFERENCE: E11PRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110
SOFTWARE: PASTESEQ For Windows Version 4.0
SEQ ID NO: 10385
LENGTH: 382
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10385

Query Match 74.4%; Score 29; DB 9; Length 382;
Bseq. Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYNAVSSI 8
Db 62 GYNAVSSI 69

Search completed: January 29, 2004, 15:01:44
Job time : 25.5441 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:44:54 : Search time 35.0735 Seconds
(without alignments)
40.730 Million cell updates/sec

Title: US-10-032-950-2
Perfect score: 39
Sequence: 1 GANXSLF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	37	94.9	21	AAV67852
2	37	94.9	23	ABP53541
3	37	94.9	23	AAV50900
4	37	94.9	23	AAV67859
5	34	87.2	9	AAV67860
6	34	87.2	9	AAV50907
7	34	87.2	9	AAV51002
8	34	87.2	9	AAV51003
9	34	87.2	23	AAV51004

10	34	87.2	9	23	ABP07160	Peptide-mediated q
11	33	84.6	9	21	AAV67851	S. aureus peptide
12	33	84.6	9	21	AAV67851	Staphylococcus aur
13	33	84.6	9	23	ABP53540	Cyclic peptide SBQ
14	33	84.6	9	23	AAV50899	Agd2-autoinducing
15	33	84.6	9	23	AAV50908	Protected peptide
16	33	84.6	9	23	AAV51005	Agd2 lactam cycli
17	33	84.6	9	23	ABP07161	Peptide-mediated q
18	32	82.1	9	19	AAV38333	Transcription inh
19	32	82.1	9	23	AAV51001	Agd2 thiolacton
20	32	82.1	9	23	ABP07159	Peptide-mediated q
21	32	82.1	9	24	ABP84631	S. aureus RN6607 a
22	32	82.1	47	24	ABP84634	S. aureus RN6607 A
23	32	82.1	71	21	AAV60552	Arabidopsis thalia
24	32	82.1	72	21	AAV6185	Arabidopsis thalia
25	32	82.1	78	21	AAV61267	Arabidopsis thalia
26	32	82.1	105	21	AAV61255	Arabidopsis thalia
27	32	82.1	112	21	AAV61255	Arabidopsis thalia
28	32	82.1	1358	23	ABP80603	Human sbg1018172CS
29	32	82.1	3105	23	ABP80604	Human sbg1018172CS
30	31	79.5	536	23	ABP93222	Heridically activ
31	30	76.9	9	21	AAV67855	S. aureus peptide
32	30	76.9	9	21	AAV67856	S. aureus peptide
33	30	76.9	9	23	ABP53544	Cyclic peptide SBQ
34	30	76.9	9	23	ABP53545	Agd2-autoinducing
35	30	76.9	9	23	AAV50903	Novel human transp
36	30	76.9	258	21	AAV50904	Novel human transp
37	30	76.9	258	21	AAV50904	Novel human transp
38	30	76.9	258	21	AAV50904	Novel human transp
39	30	76.9	672	21	AAV50904	Novel human transp
40	30	76.9	672	21	AAV50904	Novel human transp
41	30	76.9	771	23	ABP53544	Human polypeptide
42	30	76.9	771	24	AAV53544	Human transporter
43	30	76.9	895	23	ABP53544	Heridically activ
44	30	76.9	1570	22	ABP71314	Drosophila melanog
45	30	76.9	1570	22	ABP71316	Drosophila melanog

ALIGNMENTS

RESULT 1	AAV67852	standard; peptide; 9 AA.
ID	AAV67852	
XX	AAV67852	
AC	AAV67852	
XX	AAV67852	
DT	25-APR-2000	(first entry)
XX	25-APR-2000	
DE	S. aureus peptide #2 used for bacterial interference.	
KW	Staphylococcus aureus infection; cyclic peptide; Agd2; agr response;	
KM	virulence factor; treatment.	
XX	Staphylococcus aureus.	
OS	Staphylococcus aureus.	
XX	Staphylococcus aureus.	
FT	Key	Location/Qualifiers
FT	Modified-site	1
FT	/note= "N-terminal residue forms bond with C-terminal residue to form a cyclic peptide"	
FT	Misc-difference	5
FT	/label= Unknown	
XX	W0967286-A2.	
XX	29-DEC-1999.	
PD	29-DEC-1999.	
XX	24-JUN-1999;	99WO-US14562.
PF	24-JUN-1999;	99US-0103438.
XX	24-JUN-1999;	99US-0103438.
PR	24-JUN-1999;	99US-0103438.
XX	(UVR0) UNIV ROCKEFELLER.	
XX	(UVR0) UNIV NEW YORK STATE.	
PA	(UVR0) UNIV NEW YORK STATE.	

```

XX  PT Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
XX  DR WPI; 2000-147202/13.
XX  CC New cyclic peptides for treating infections with Staphylococcus aureus
XX  PT
XX  PS Claim 9; Page 26; 37pp; English.
XX  CC This sequence represents a cyclic peptide derived from the Staphylococcus
XX  CC aureus AgrD peptide. The invention relates to AgrD derived peptides,
XX  CC a composition containing a peptide and a carrier, and a method for the
XX  CC production of the cyclic peptides. The peptide inhibits the agr response,
XX  CC which is normally associated with the release of virulence factors of
XX  CC Staphylococcus aureus. An AgrD peptide is produced by S. aureus that
XX  CC activates the agr response in strains of a single group, but interferes
XX  CC with this response in strains of different groups. The peptides and
XX  CC composition containing them can be used to treat infections by S. aureus.
XX  SQ Sequence 9 AA;
XX  Query Match 94.9%; Score 37; DB 21; Length 9;
XX  Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GANAXSSLF 9
   |||||
Db 1 GANAXSSLF 9

RESULT 2
ABP53541
ID ABP53541 standard; peptide; 9 AA.
XX AC ABP53541;
XX DT 13-DEC-2002 (first entry)
XX DE Cyclic peptide SEQ ID NO:2.
XX KM Cyclic peptide; Staphylococcus aureus; infection; antibacterial;
XX KM agr response inhibitor.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 5 /note= "any amino acid"
XX PN US200207453-A1.
XX PD 20-JUN-2002.
XX PF 27-DEC-2001; 2001US-0032950.
XX PR 24-JUN-1998; 98US-090402P.
XX PR 24-JUN-1999; 99US-0339511.
XX PA (MUIR/) MUJR T W.
XX PA (MAYV/) MAYVILLE P.
XX PA (NOVI/) NOVICK R P.
XX PA (BEAV/) BEAVIS R.
XX PA (JIG/) JI G.
XX PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX PS WPI; 2002-681366/73.
XX DR WPI; 2002-681366/73.
XX PT New cyclic peptides, useful for treating Staphylococcus aureus
XX PT infections
XX PS Claim 9; Page 10; 18pp; English.

```

```

XX  CC ABP53540 to ABP53547 represent cyclic peptides (I) from the present
XX  CC invention. The present invention also describes a method for treating
XX  CC Staphylococcus aureus infection comprising the administration of a
XX  CC composition comprising (I). (I) has antibacterial activity, and can be
XX  CC used as an agr gene response inhibitor. The peptides are useful for
XX  CC treating S. aureus infections.
XX  SQ Sequence 9 AA;
XX  Query Match 94.9%; Score 37; DB 23; Length 9;
XX  Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GANAXSSLF 9
   |||||
Db 1 GANAXSSLF 9

RESULT 3
AAM50900
ID AAM50900 standard; peptide; 9 AA.
XX AC AAM50900;
XX DT 08-MAY-2002 (first entry)
XX DE AgrD-autoinducing cyclic peptide, inhibitor of agr response.
XX KM Staphylococcus aureus; AgrD; agr response; inhibitor; antibiotic;
XX KM antibacterial; infection; therapy; cyclic.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 5 /note= "any amino acid"
XX FT Misc-difference 5 /note= "note linked to residue 9 to form cyclic
XX FT peptide"
XX FT Misc-difference 9 /note= "note linked to residue 5 to form cyclic
XX FT peptide"
XX PN US637385-B1.
XX PD 08-JAN-2002.
XX PF 24-JUN-1999; 99US-0339511.
XX PR 24-JUN-1998; 98US-090402P.
XX PA (UYRO ) UNIV ROCKEFELLER.
XX PA (UYNY ) UNIV NEW YORK STATE.
XX PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX PS WPI; 2002-170774/22.
XX PT Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
XX PT interference and for treating Staphylococcus aureus infection in a
XX PT subject
XX PS Claim 7; Column 19; 18pp; English.
XX CC The present sequence is that of a novel synthetic cyclic peptide
XX CC of the invention that is capable of inhibiting the agr response of
XX CC Staphylococcus aureus. It is an AgrD-autoinducing peptide, where
XX CC AgrD is a secreted agr-encoded peptide and where the agr locus
XX CC controls the synthesis of virulence factor and other extracellular
XX CC proteins responsible for pathogenicity in S. aureus. Preferred
XX CC peptides may have the sequence NH2-X(n)-Z-X(y)-COOH, with a
XX CC cyclic bond between the Z residue and COOH other than a thioester

```

CC bond, where X is an amino acid, an amino acid analogue, a
 CC peptidomimetic or non-amide isostere, Z is a synthetic or a
 CC biosynthetic amino acid, n is 0-10 and Y is 1-10. The cyclic bond
 CC is especially a lactam or lactone bond. The thiololactone
 CC structure within native AgrD peptides is required for activation
 CC of the agr response. Elimination of the thiol ester component of
 CC the cyclic ring structure can destroy agr response activating.
 CC activity while preserving and enhancing inhibitory activity. A
 CC claimed method of preparing a cyclic peptide involves: assembling
 CC a linear peptide chain on to a solid phase resin support;
 CC deprotecting the resulting protected assembled peptide; treating the
 CC deprotected peptide with neutral buffer for a time sufficient to
 CC form the cyclic peptide and cleave the peptide from the support;
 CC and recovering the cyclic peptide. The peptide is useful for
 CC bacterial interference, especially for the treatment of S. aureus
 CC infection.

XX Sequence 9 AA;

Query Match 94.9%; Score 37; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
 1 GANAXSLF 9

Db

RESULT 4
 AAY67859 standard; peptide; 9 AA.

XX AAY67859;

XX 25-APR-2000 (first entry)

XX Staphylococcus aureus AgrDII derived peptide sequence.

XX Staphylococcus aureus infection; treatment; AgrD; agr response;

XX virulence factor.

XX Staphylococcus aureus.

XX WO967286-A2.

XX 29-DEC-1999.

XX 24-JUN-1999; 99WO-US14562.

XX 24-JUN-1998; 98US-0103438.

XX (UYRQ) UNIV ROCKEFELLER.

XX (UYNY) UNIV NEW YORK STATE.

XX Muir TW, Mayville P, Novick RP, Ji G, Beavis R;

XX WPI; 2000-147202/13.

XX New cyclic peptides for treating infections with Staphylococcus aureus

XX Examples; Page 22; 37pp; English.

CC This sequence represents the Staphylococcus aureus AgrDII derived
 CC peptide. The invention relates to AgrD derived cyclic peptides, a
 CC composition containing a peptide and a carrier, and a method for the
 CC production of the cyclic peptides. The peptide inhibits the agr response,
 CC which is normally associated with the release of virulence factors of
 CC Staphylococcus aureus. An AgrD peptide is produced by S. aureus that
 CC activates the agr response in strains of a single group, but interferes
 CC with this response in strains of different groups. The peptides and
 CC composition containing them can be used to treat infections by S. aureus.

SQ Sequence 9 AA;

Query Match 87.2%; Score 34; DB 21; Length 9;
 Best Local Similarity 77.8%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
 1 GANAXSLF 9

Db

RESULT 5
 AAY67860 standard; peptide; 9 AA.

XX AAY67860;

XX 25-APR-2000 (first entry)

XX Staphylococcus aureus AgrDII derived peptide sequence.

XX Staphylococcus aureus infection; AgrD; agr response; treatment;

XX virulence factor.

XX Staphylococcus aureus.

XX WO967286-A2.

XX 29-DEC-1999.

XX 24-JUN-1999; 99WO-US14562.

XX 24-JUN-1998; 98US-0103438.

XX (UYRQ) UNIV ROCKEFELLER.

XX (UYNY) UNIV NEW YORK STATE.

XX Muir TW, Mayville P, Novick RP, Ji G, Beavis R;

XX WPI; 2000-147202/13.

XX New cyclic peptides for treating infections with Staphylococcus aureus

XX Examples; Page 22; 37pp; English.

CC This sequence represents the Staphylococcus aureus AgrDII derived
 CC peptide. The invention relates to AgrD derived cyclic peptides, a
 CC composition containing a peptide and a carrier, and a method for the
 CC production of the cyclic peptides. The peptide inhibits the agr response,
 CC which is normally associated with the release of virulence factors of
 CC Staphylococcus aureus. An AgrD peptide is produced by S. aureus that
 CC activates the agr response in strains of a single group, but interferes
 CC with this response in strains of different groups. The peptides and
 CC composition containing them can be used to treat infections by S. aureus.

SQ Sequence 9 AA;

Query Match 87.2%; Score 34; DB 21; Length 9;
 Best Local Similarity 77.8%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
 1 GANAXSLF 9

Db

RESULT 6
 AAM50907 standard; Peptide; 9 AA.

XX AAM50907;

JT		08-MAY-2002 (first entry)	
DE	Protected peptide used in cyclic peptide production.		
XW	Staphylococcus aureus; AgrD; agr response; inhibitor; antibiotic;		
KX	antibacterial; infection; therapy; cyclic.		
XX	Synthetic.		
FH	Key	Location/Qualifiers	
FT	Modified-site 1	/note= "Z-gly"	
FT	Modified-site 5	/note= "Ser(Tbu)"	
FT	Modified-site 6	/note= "ser(BzI) "	
FT	Modified-site 7	/note= "Ser(Bzl) "	
FT	Misc-difference 5	/note= "note linked to residue 9 to form cyclic peptide"	
FT	Misc-difference 9	/note= "note linked to residue 5 to form cyclic peptide"	
PN	US6337385-B1.		
PD	08-JAN-2002.		
PF	24-JUN-1999; 99US-0339511.		
PR	24-JUN-1998; 98US-090402P.		
PA	(UYRQ) UNIV ROCKEFELLER. (UYNV) UNIV NEW YORK STATE.		
PI	Muir TM, Mayville P, Novick RP, Beavis R, Ji G; MPI; 2002-170774/22 .		
PT	Novel synthetic, cyclic AgRP-autoinducing peptide for bacterial interference and for treating Staphylococcus aureus infection in a subject - Disclosure; Column 14; 16pp; English.		
CC	The present sequence is that of a protected peptide used in an example of the preparation of novel synthetic cyclic peptides of the invention (see AIMS0899-906). The peptide corresponds to the Staphylococcus aureus AgrdII sequence with a Cys5 to Ser mutation (lactone). It was synthesised on a Wang-resin using an Fmoc N-alpha protection strategy. Following chain assembly, the peptide was cleaved from the support and the ser-5 residue deprotected by treatment with a trifluoroacetic acid/anisole/water mixture (90:5:5), for 4 hr. The partially protected peptide-alpha carbonyls were then dissolved in DMF and treated with PyBOP and a catalytic amount of diethylaminopyridine. Cyclization was complete after 2 hr. The remaining protecting groups were then removed by treatment with HF and the peptide purified by HPLC. The cyclic peptide is capable of inhibiting the agr response of Staphylococcus aureus. The thiolactone structure within native AgrD peptides is required for activation of this response. Replacement of the thiol ester component of the cyclic ring structure with a lactone (as in the present case) or a lactam can destroy agr response activating activity while preserving and enhancing inhibitory activity. The cyclic peptides are useful for bacterial interference, especially for the treatment of S. aureus infection.		
SC	Sequence 9 AA:		
Query Match	87.2%; Score 34; DB 23; Length 9;		
.Best Local Similarity	77.8%; Pred. No. 9.3e+05;		

Matches	7;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
QY	1	GANAXSLF 9							
Db	1	GVMASSSLF 9							
<hr/>									
RESULT 7									
ID	AAM51002	AAM51002 standard; Peptide; 9 AA.							
XX	AC	AAM51002;							
XX	DT	08-MAY-2002 (first entry)							
XX	DE	AgrD2 linear thioester peptide.							
XX	KM	AgrD2; agr response; inhibitor; antibiotic; antibacterial;							
XX	KW	Infection; therapy.							
XX	OS	Staphylococcus aureus.							
XX	OS	Synthetic.							
FH	Key	Location/Qualifiers							
FT	Modified-site	9 /note= "C-terminal thioester"							
PN	US6373785-B1.								
XX	PD	08-JAN-2002.							
XX	PP	24-JUN-1999; 99US-0339511.							
XX	PR	24-JUN-1998; 98US-090402P.							
XX	PA	(UYRO) UNIV ROCKEFELLER.							
XX	PA	(UYNV) UNIV NEW YORK STATE.							
P1	Maier TW,	Mayville P,	Novack RP,	Beavis R,	Ji G;				
DR	WPI;	2002-170774/22.							
XX	PT	Novel synthetic, cyclic AgrD-anticoagulating peptide for bacterial							
PT	interference	and for treating Staphylococcus aureus infection in a							
PT	subject	-							
XX	Example 1;	Column 9; 18pp; English.							
XX	PS	The present sequence is that of a novel synthetic AgrD2 linear							
CC	thioester peptide.	The peptide is derived from the cyclic AgrD2							
CC	peptide of Staphylococcus aureus group II. AgrD2 is a secreted								
CC	agr-encoded peptide, where the agr locus controls the synthesis of								
CC	virulence factor and other extracellular proteins responsible for								
CC	pathogenicity in S. aureus. The biological activity of the								
CC	synthetic peptide was assayed using cultured S. aureus strains								
CC	containing a beta-lactamase reporter gene fused to the agrP3								
CC	promoter. This allowed activation or inhibition of the agr								
CC	response to be monitored spectrophotometrically. Unlike an AgrD2								
CC	ethanolactone cyclic peptide (see AAM51001), the present peptide								
CC	was unable to either activate or inhibit the agr response, even								
CC	when added to cultured cells at uM concentrations. The invention								
CC	provides claimed cyclic peptides (see AAM50899-906 and AAM50999)								
CC	and methods for preparing them. The cyclic peptides are useful for								
CC	bacterial interference, especially for the treatment of S. aureus								
CC	infection.								
XX	Sequence	9 AA;							
XX	Query Match	87.2%; Score 34; DB 23; Length 9;							
XX	Best Local Similarity	77.8%; Pred.No. 9.3e+05;							
XX	Matches	7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;							
QY	1	GANAXSLF 9							

Db 1 GVNAASSLF 9

RESULT 8

AAMS1003
ID AAMS1003 standard; Peptide; 9 AA.
XX
XX AAMS1003;

DT 08-MAY-2002 (first entry)

DE AgRD2 linear free acid peptide.

KM AgRD2; agr response; inhibitor; antibiotic; antibacterial;
infection; therapy.

OS Staphylococcus aureus.
Synthetic.

US6337385-B1.

08-JAN-2002.

24-JUN-1999; 99US-0339511.

24-JUN-1998; 98US-090402P.

(UYRQ) UNIV ROCKEFELLER.

(UYNV) UNIV NEW YORK STATE.

Muir TW, Mayville P, Novick RP, Beavis R, Ji G;

WPI; 2002-170774/22.

Novel synthetic, cyclic AgRD2-autoinducing peptide for bacterial interference and for treating Staphylococcus aureus infection in a subject

Example 1; Column 9; 18bp; English.

The present sequence is that of a novel synthetic AgRD2 linear free acid peptide. The peptide is derived from the cyclic AgRD2 peptide of Staphylococcus aureus group II. AgRD2 is a secreted agr-encoded peptide, where the agr locus controls the synthesis of virulence factor and other extracellular proteins responsible for pathogenicity in S. aureus. The biological activity of the synthetic peptide was assayed using cultured S. aureus strains containing a beta-lactamase reporter gene fused to the agr3 promoter. This allowed activation or inhibition of the agr response to be monitored spectrophotometrically. Unlike an AgRD2 thiolactone cyclic peptide (see AAMS1001), the present peptide was unable to either activate or inhibit the agr response, even when added to cultured cells at low concentrations. The invention provides claimed cyclic peptides (see AAMS0899-906 and AAMS0999) and methods for preparing them. The cyclic peptides are useful for bacterial interference, especially for the treatment of S. aureus infection.

Sequence 9 AA;

Query Match 87.2%; Score 34; DB 23; Length 9;
Best Local Similarity 77.8%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANAXSSLF 9
Db 1 GVNAASSLF 9

RESULT 9
AAMS1004
ID AAMS1004 standard; Peptide; 9 AA.

XX AAMS1004;

DT 08-MAY-2002 (first entry)

DE AgRD2 lactone cyclic peptide.

KM AgRD2; agr response; inhibitor; antibiotic; antibacterial;
infection; therapy; lactone; cyclic.

OS Staphylococcus aureus.
Synthetic.

US6337385-B1.

08-JAN-2002.

24-JUN-1999; 99US-0339511.

24-JUN-1998; 98US-090402P.

(UYRQ) UNIV ROCKEFELLER.

(UYNV) UNIV NEW YORK STATE.

Muir TW, Mayville P, Novick RP, Beavis R, Ji G;

WPI; 2002-170774/22.

Novel synthetic, cyclic AgRD2-autoinducing peptide for bacterial interference and for treating Staphylococcus aureus infection in a subject

Example 1; Column 9; 18bp; English.

The present sequence is that of a novel synthetic AgRD2 lactone cyclic peptide in which residue 5 of the peptide is linked to residue 9 via a lactone bond. The peptide is derived from an AgRD2 peptide of Staphylococcus aureus group II. AgRD2 is a secreted agr-encoded peptide, where the agr locus controls the synthesis of virulence factor and other extracellular proteins responsible for pathogenicity in S. aureus. The biological activity of the synthetic peptide was assayed using cultured S. aureus strains containing a beta-lactamase reporter gene fused to the agr3 promoter. This allowed activation or inhibition of the agr response to be monitored spectrophotometrically. The lactone AgRD2 peptide inhibited the agr response of group I S. aureus strains without activating the agr response in group I, II or III strains. The invention provides claimed cyclic peptides (see AAMS0899-906 and AAMS0999) and methods for preparing them, especially peptides where the cyclic bond is a lactam or lactone bond. The cyclic peptides are useful for bacterial interference, especially for the treatment of S. aureus infection.

Sequence 9 AA;

Query Match 87.2%; Score 34; DB 23; Length 9;
Best Local Similarity 77.8%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANAXSSLF 9
Db 1 GVNAASSLF 9

RESULT 10

ID	ABB07160	standard; peptide; 9 AA.
AC	ABB07160;	
XX		
DT	13-MAR-2002	(first entry)
XX		
DE	Peptide-mediated quorum sensing inhibitor peptide cyclo-XII.	
XX		
KM	Autoinducer-2; AI-2; antibiotic; antibacterial; dermatological; -	
KM	vulnerable; pheromone; agr system; accessory gene regulator; cyclic.	
XX		
OS	Synthetic.	
XX		
PN	WO200185664-A2.	
PD	15-NOV-2001.	
XX		
PF	10-MAY-2001; 2001WO-US15621.	
PR	10-MAY-2000; 2000US-203000P.	
PR	07-DEC-2000; 2000US-254398P.	
XX		
PA	(UYPR-) UNIV PRINCETON.	
PA	(CUOR-) CUOREX PHARM INC.	
PA	(UYTE-) UNIV TECHNOLOGIES INT INC.	
PI	Basel BL, Dammel CS, Schauder S, Shokat K, Stein J, Surette MG;	
DR	WPI; 2002-075235/10.	
XX		
PT	Use of autoinducer-2 agonists or antagonists for regulating activity of	
PT	autoinducer-2 receptor, regulating bacterial growth and pathogenesis,	
XX	also antibiotic compositions -	
XX		
BS	Disclosure; Page 33; 134pp; English.	
CC	The invention relates to the use of autoinducer-2 (AI-2) agonists or	
CC	antagonists for regulating activity of autoinducer-2 receptor, regulating	
CC	bacterial growth and pathogenesis. Synergistic antibiotic compositions	
CC	comprising inhibitors of the quorum-sensing pathway of a microorganism	
CC	are also provided. Methods using such AI-2 analogues are useful for	
CC	treating pathogen-associated disease states. The compounds and antibiotic	
CC	compositions can be used to inhibit bacterial cell growth and/or biofilm	
CC	formation on a medical device, particularly for promoting growth of skin	
CC	graft replacements used in the treatment of burns and ulcers. They may	
CC	also be used to aid wound repair, and to inhibit bacterial cell growth	
CC	and biofilm formation in or on products or devices used for personal	
CC	hygiene. The present sequence represents a inhibitor of peptide-mediated	
CC	quorum sensing.	
XX		
XX		
SQ	Sequence 9 AA;	
	Query Match 87.2%; Score 34; DB 23; Length 9;	
	Best Local Similarity 77.8%; Prid. No. 9,3e+05;	
	Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0	
OY	1 GANXSSLP 9	
DB	1 GVNASSLP 9	
RSRULT 11		
ID	AAV67851	
XX	AAV67851 standard; peptide; 9 AA.	
XX		
XX	AAV67851;	
DT	25-APR-2000	(first entry)
XX		
DE	S. aureus peptide #1 used for bacterial interference.	
KM	Staphylococcus aureus infection; cyclic peptide; AgrD; agr response;	

```

KW virulence factor; treatment.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Misc-difference 5 /label= Unknown
XX
XX
XX W0967286-A2.
XX
XX 29-DEC-1999.
XX
XX 24-JUN-1999; 99WO-US14562.
XX
XX 24-JUN-1998; 98US-0103438.
XX
XX (UYRQ ) UNIV ROCKEFELLER.
XX (UYNY ) UNIV NEW YORK STATE.
XX
XX Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
XX WPI; 2000-147202/13.
XX
XX New cyclic peptides for treating infections with Staphylococcus aureus
XX
XX Claim 9; Page 26; 37pp; English.
XX
XX This sequence represents a cyclic peptide derived from the Staphylococcus
XX aureus AgrD peptide. The invention relates to AgrD derived peptides,
XX a composition containing a peptide and a carrier, and a method for the
XX production of the cyclic peptides. The peptide inhibits the agr response,
XX which is normally associated with the release of virulence factors of
XX Staphylococcus aureus. An AgrD peptide is produced by S. aureus that
XX activates the agr response in strains of a single group, but interferes
XX with this response in strains of different groups. The peptides and
XX composition containing them can be used to treat infections by S. aureus.
XX
XX
XX Sequence 9 AA;
SQ
Query Match 84.6%; Score 33; DB 21; Length 9;
Best Local Similarity 88.9%; Pred. No. 9.3e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0.
QY 1 GANAXSLF 9
Db 1 GANAXSLF 9
RESULT 12
AAV67861
ID AAV67861 standard; peptide; 9 AA.
XX
XX AAV67861;
XX
XX 25-APR-2000 (first entry)
XX
XX Staphylococcus aureus AgrDII derived peptide sequence.
XX
XX Staphylococcus aureus infection; AgrD; agr response; treatment;
XX virulence factor.
XX
XX Staphylococcus aureus.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 5 /label= Unknown
XX
XX W0967286-A2.
XX
XX 29-DEC-1999.
XX
XX 24-JUN-1999; 99WO-US14562.
XX

```

XX 24-JUN-1998; 98US-0103438.
 XX (UYRQ) UNIV ROCKEFELLER.
 PA (UYNY) UNIV NEW YORK STATE.
 XX Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
 XX WPI; 2000-147202/13.
 XX New cyclic peptides for treating infections with Staphylococcus aureus
 PT
 PS Examples; Page 22, 37pp; English.
 XX This sequence represents the Staphylococcus aureus AgrDII derived
 CC peptide. The invention relates to Agrd derived cyclic peptides, a
 CC composition containing a peptide and a carrier, and a method for the
 CC production of the cyclic peptides. The peptide inhibits the agr response,
 CC which is normally associated with the release of virulence factors of
 CC Staphylococcus aureus. An Agrd peptide is produced by S. aureus that
 CC activates the agr response in strains of a single group, but interferes
 CC with this response in strains of different groups. The peptides and
 CC composition containing them can be used to treat infections by S. aureus.
 XX
 SQ Sequence 9 AA;
 Query Match 84.6%; Score 33; DB 21; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GANAXSSLP 9
 Db 1 GVNAXSSLP 9
 RESULT 13
 ABP53540 ID ABP53540 standard; peptide; 9 AA.
 AC ABP53540;
 XX 13-DEC-2002 (first entry)
 DT
 XX Cyclic peptide SEQ ID NO:1.
 DE Cyclic peptide; Staphylococcus aureus; infection; antibacterial;
 KM agr response inhibitor.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 5 /note= "any amino acid"
 FT
 XX US2002077453-A1.
 FN
 XX 20-JUN-2002.
 PD
 XX 27-DEC-2001; 2001US-0032950.
 PF
 XX 24-JUN-1998; 98US-090402P.
 FR 24-JUN-1999; 99US-0339511.
 XX (MUIR/) MUIR T W.
 PA (MAYV/) MAYVILLE P.
 PA (NOVI/) NOVICK R P.
 PA (BEAV/) BEAVIS R.
 PA (JIGG/) JI G.
 XX Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
 PI WPI; 2002-681366/73.

XX New cyclic peptides, useful for treating Staphylococcus aureus
 PT infections -
 PT
 XX Claim 9; Page 10; 18pp; English.
 PS
 XX ABP53540 to ABP53547 represent cyclic peptides (I) from the present
 CC invention. The present invention also describes a method for treating
 CC Staphylococcus aureus infection comprising the administration of a
 CC composition comprising (I). (I) has antibacterial activity, and can be
 CC used as an agr gene response inhibitor. The peptides are useful for
 CC treating S. aureus infections.
 XX
 SQ Sequence 9 AA;
 Query Match 84.6%; Score 33; DB 23; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GANAXSSLP 9
 Db 1 GVNAXSSLP 9
 RESULT 14
 AAM50899 ID AAM50899 standard; peptide; 9 AA.
 AC AAM50899;
 XX 08-MAY-2002 (first entry)
 DT
 XX Agrd-autoinducing cyclic peptide, inhibitor of agr response.
 DE
 XX Staphylococcus aureus; Agrd; agr response; inhibitor; antibiotic;
 KM antibacterial; infection; therapy; cyclic.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 5 /note= "any amino acid"
 FT
 XX Misc-difference 5 /note= "note linked to residue 9 to form cyclic
 FT peptide"
 FT Misc-difference 9 /note= "note linked to residue 5 to form cyclic
 FT peptide"
 XX US6337385-B1.
 FN
 XX 08-JAN-2002.
 PD
 XX 24-JUN-1999; 99US-0339511.
 PF
 XX 24-JUN-1998; 98US-090402P.
 FR (UYRQ) UNIV ROCKEFELLER.
 PA (UYNY) UNIV NEW YORK STATE.
 XX Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
 PI WPI; 2002-170774/22.
 XX Novel synthetic, cyclic Agrd-autoinducing peptide for bacterial
 PT interference and for treating Staphylococcus aureus infection in a
 PT subject -
 PT
 PS Claim 7; Column 19; 18pp; English.
 XX The present sequence is that of a novel synthetic cyclic peptide
 CC of the invention that is capable of inhibiting the agr response of
 CC Staphylococcus aureus. It is an Agrd-autoinducing peptide, where

CC AgrD is a secreted agr-encoded peptide and where the agr locus
 CC controls the synthesis of virulence factor and other extracellular
 CC proteins responsible for pathogenicity in *S. aureus*. Preferred
 CC peptides may have the sequence NH₂-X(n)-Z-X(y)-COOH, with a
 CC cyclic bond between the Z residue and COOH other than a thioester
 CC bond, where X is an amino acid, an amino acid analogue, a
 CC peptidomimetic or non-amide isostere, Z is a synthetic or a
 CC biosynthetic amino acid, n is 0-10 and y is 1-10. The cyclic bond
 CC is especially a lactam or lactone bond. The thiololactone
 CC structure within native AgrD peptides is required for activation
 CC of the agr response. Elimination of the thiol ester component of
 CC the cyclic ring structure can destroy agr response activating
 CC activity while preserving and enhancing inhibitory activity. A
 CC claimed method of preparing a cyclic peptide involves: assembling
 CC a linear peptide chain on to a solid phase resin support;
 CC deprotecting the resulting protected assembled peptide; treating the
 CC deprotected peptide with neutral buffer for a time sufficient to
 CC form the cyclic peptide and cleave the peptide from the support;
 CC and recovering the cyclic peptide. The peptide is useful for
 CC bacterial interference, especially for the treatment of *S. aureus*
 CC infection.

CC Sequence 9 AA;

Query Match 84.6%; Score 33; DB 23; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANXXSLF 9
 | | | | |
 DB 1 GANXXSLF 9

RESULT 15

AAAM50908
 ID AAAM50908 standard; Peptide; 9 AA.

AC AAAM50908;

DT 08-MAY-2002 (first entry)

DE Protected peptide used in cyclic peptide production.

KW *Staphylococcus aureus*; AgrD; agr response; inhibitor; antibiotic;

KW antibacterial; infection; therapy; cyclic.

OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "Z-Gly"

FT Modified-site 5 /label= Dpr (Boc)

FT Modified-site 6 /note= "Ser(Bzl)"

FT Modified-site 7 /note= "Ser(Bzl)"

FT Misc-difference 5 /note= "note linked to residue 9 to form cyclic peptide"

FT Misc-difference 9 /note= "note linked to residue 5 to form cyclic peptide"

FT

PN US637385-B1.

PD 08-JAN-2002.

PF 24-JUN-1999; 99US-0339511.

PR 24-JUN-1998; 98US-090402P.

PA (UYRQ) UNIV ROCKEFELLER.

PA (UYNY) UNIV NEW YORK STATE.
 XX
 XX Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
 XX WPI; 2002-170774/22.

DR Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
 FT interference and for treating *Staphylococcus aureus* infection in a
 PT subject -
 XX Disclosure; Column 14; 18pp; English.

XX The present sequence is that of a protected peptide used in an
 CC example of the preparation of novel synthetic cyclic peptides of
 CC the invention (see AAAM50899-906). The peptide corresponds to the
 CC *Staphylococcus aureus* AgrDII sequence with a Cys5 to diaminopropionic
 CC acid (Dpr) mutation (lactam). It was synthesised on a Wang-resin
 CC using an Fmoc N-alpha protection strategy. Following chain assembly,
 CC the peptide was cleaved from the support and the Dpr-5 residue
 CC deprotected by treatment with a trifluoroacetic acid:anisole:water
 CC mixture (90:5:5) for 4 hr. The partially protected peptide-alpha
 CC carboxylates were then dissolved in DMF and treated with PyBOP.
 CC Cyclization was complete after 2 hr. The remaining protecting groups
 CC were removed by treatment with HF and the peptide purified by HPLC.
 CC The cyclic peptide is capable of inhibiting the agr response of
 CC *Staphylococcus aureus*. The thiololactone structure within native
 CC AgrD peptides is required for activation of this response.
 CC Replacement of the thiol ester component of the cyclic ring
 CC structure with a lactam (as in the present case) or a lactone can
 CC destroy agr response activating activity while preserving and
 CC enhancing inhibitory activity. The cyclic peptides are useful for
 CC bacterial interference, especially for the treatment of *S. aureus*
 CC infection.

XX Sequence 9 AA;

Query Match 84.6%; Score 33; DB 23; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANXXSLF 9
 | | | | |
 DB 1 GANXXSLF 9

Search completed: January 29, 2004, 14:52:10
 Job time : 35.0735 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:46:49 ; Search time 11.1176 Seconds
(without alignments)
77.851 Million cell updates/sec

Title: US-10-032-950-2

Perfect score: 39

Sequence: 1 GANAXSLF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR:76:*
2: PIR:1:*
3: PIR:3:*
4: PIR:4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	84.6	269	2	SS8439
2	33	84.6	382	2	A28067
3	33	84.6	405	2	A60534
4	33	84.6	407	2	A30200
5	32	82.1	47	2	C89995
6	32	82.1	453	1	ZPBOC2
7	32	82.1	573	1	FE1313
8	31	79.5	340	2	S76294
9	31	79.5	536	2	T04998
10	31	79.5	1232	2	T06165
11	30	76.9	213	2	FE1314
12	30	76.9	250	2	H84428
13	30	76.9	291	2	C89103
14	30	76.9	333	1	C69812
15	29	76.9	895	2	T45786
16	29	74.4	313	2	E90634
17	29	74.4	313	2	E85485
18	29	74.4	313	2	E64725
19	29	74.4	342	2	B85757
20	29	74.4	457	2	A99862
21	29	74.4	476	2	D86306
22	29	74.4	551	2	C84549
23	29	74.4	870	2	S74291
24	29	74.4	1435	2	T32930
25	29	74.4	2124	2	H83357
26	28	71.8	771	2	T16158
27	28	71.8	186	2	B87290
28	28	71.8	211	2	A40899
29	28	71.8	226	2	D65105

ALIGNMENTS

30	28	71.8	226	2	A98133	hypothetical prote
31	28	71.8	226	2	D85978	hypothetical prote
32	28	71.8	344	2	S64931	hypothetical prote
33	28	71.8	362	2	A84187	hypothetical prote
34	28	71.8	420	2	AD0105	probable sugar-bin
35	28	71.8	429	2	B71307	conserved hypothet
36	28	71.8	453	2	S57097	indoleamine-pyrol
37	28	71.8	453	2	T24127	probable chitinase
38	28	71.8	486	2	G96713	probable DNA-bindi
39	28	71.8	513	2	T05210	hypothetical prote
40	28	71.8	546	2	B85398	hypothetical prote
41	28	71.8	715	2	T41402	hypothetical prote
42	28	71.8	926	2	S48463	SEC24 protein - ye
43	28	71.8	1200	2	A81343	DNA-directed DNA p
44	28	71.8	1216	2	T02146	coatomer complex a
45	28	71.8	1938	2	A37361	probable integral

RESULT 1

transcription factor AKR - chicken

C/Species: Gallus gallus (chicken)

C/Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 24-Sep-1999

C/Accession: S58439; S58440

R/Ryan, A.K.; Tejada, M.L.; May, D.L.; Dubayova, M.; Deeley, R.G.
Submitted to the EMBL Data Library, April 1995

A/Description: Isolation and characterization of the chicken homeodomain protein AKR.

A/Reference number: S58439

A/Accession: S58439

A/Molecule type: mRNA

A/Residues: 1-269 <RY>

A/Status: nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 35-94 <RY>

A/Cross-references: EMBL:U25353

A/Function:

A/Description: negative regulator of apovLDLII (major egg yolk apolipoprotein)

A/Note: expressed in liver during early embryogenesis and in non-hepatic adult tissue

C/Superfamily: Unassigned homeobox proteins; homeobox homology

C/Keywords: DNA binding; homeobox; nucleus; transcription regulation

F/36-95/Domain: homeobox homology <HOX>

Query Match

Best Local Similarity 77.8%; Pred. No. 3.6;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANAXSLF 9

DB 222 GANAXSLF 230

RESULT 2

A28067

lysosomal membrane glycoprotein LAMP-1 - mouse

C/Species: Mus musculus (house mouse)

C/Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 20-Aug-1999

C/Accession: A28067

R/Chen, J.W.; Cha, Y.; Yuxsel, K.U.; Gracy, R.W.; August, J.T.

J. Biol. Chem. 263, 8754-8758, 1988

A/Title: Isolation and sequencing of a cDNA clone encoding lysosomal membrane glycop

A/Reference number: A28067; PMID:188243732; PMID:3379044

A/Accession: A28067

A/Molecule type: mRNA

A:Residues: 1-382 <CHE>
 A:Cross-references: GB:J03881; NID:G198706; PID:AAA39411.1; PID:G293692
 A>Note: the authors translated the codon ATT for residue 1 as Ieu and CCG for residue 2
 C:Superfamily: lysosome-associated membrane protein
 C:Keywords: glycoprotein; membrane protein

Query Match 84.6%; Score 33; DB 2; Length 382;
 Best Local Similarity 77.8%; Pred. No. 5.3;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
 DB 256 GMANSSSLF 264

RESULT 3

P2B/LAMP-1 precursor - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 03-May-1996
 C:Accession: A60534
 R:Helferman, W.; Yousefi, S.; Dennis, J.W.
 Cancer Res. 49, 6077-6084, 1989
 A>Title: Molecular characterization of P2B/LAMP-1, a major protein target of a metastasizing melanoma cell line
 A:Reference number: A60534; MUID:90002989; PMID:2676155
 A:Accession: A60534
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-405 <HEF>
 C:Superfamily: lysosome-associated membrane protein

Query Match 84.6%; Score 33; DB 2; Length 405;
 Best Local Similarity 77.8%; Pred. No. 5.6;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
 DB 279 GMANSSSLF 287

RESULT 4

120K lysosomal membrane glycoprotein precursor - rat
 N:Alternate names: sialoglycoprotein
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 18-Oct-1989 #sequence_revision 30-Sep-1991 #text_change 20-Aug-1999
 C:Accession: A30200; S03331
 R:Howe, C.L.; Granger, B.L.; Hult, M.; Green, S.A.; Gabel, C.A.; Helentius, A.; Mellman, A.; Proc. Natl. Acad. Sci. U.S.A. 85, 7577-7581, 1988
 A>Title: Derived protein sequence, oligosaccharides, and membrane insertion of the 120-K glycoprotein.
 A:Reference number: A30200; MUID:89017240; PMID:3174652
 A:Accession: A30200
 A:Molecule type: mRNA
 A:Residues: 1-407 <HOM>
 A>Note: the authors translated the codon GGG for residue 15 as Val
 R:Homeno, M.; Noguchi, Y.; Saseki, H.; Tanaka, Y.; Furuno, K.; Kono, A.; Sakaki, Y.; Kat
 FEBS Lett. 244, 351-356, 1989
 A>Title: Isolation and sequencing of a cDNA clone encoding 107 kDa sialoglycoprotein in
 A:Reference number: S03331; MUID:89153580; PMID:2920835
 A:Accession: S03331
 A:Molecule type: mRNA
 A:Residues: 22-407 <HIM>
 A:Cross-references: EMBL:X14765; NID:G56577; PID:CA432873.1; PID:G56578
 A>Note: part of this sequence, including the amino end of the mature protein, was confir
 C:Superfamily: lysosome-associated membrane protein
 C:Keywords: glycoprotein; membrane protein
 F:22-407/Product: lysosomal membrane glycoprotein, 107K #status experimental <MAT>

Query Match 84.6%; Score 33; DB 2; Length 407;
 Best Local Similarity 77.8%; Pred. No. 5.6;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
 DB 281 GMANSSSLF 289

RESULT 5

Agd protein [imported] - Staphylococcus aureus (strain N315)
 C:Species: Staphylococcus aureus
 C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:Accession: C89995
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; C
 ma, A.; Mizutani, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, I
 C.; Shiba, T.; Hattori, W.; Ogasawara, N.; Hayashi, H.; Hiratake, K.
 Lancet 357, 1225-1240, 2001
 A>Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A:Reference number: A89758; MUID:21311952; PMID:11418146
 A:Accession: C89995
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-47 <KUD>
 A:Cross-references: GB:BA000018; PID:G13701831; PID:BAE43124.1; GSPDB:GN00149
 A:Experimental source: strain N315
 A:Genetics:

Query Match 82.1%; Score 32; DB 2; Length 47;
 Best Local Similarity 77.8%; Pred. No. 1;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
 DB 24 GMANSSSLF 32

RESULT 6

ZBPOC
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) core protein II precursor - bovine
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 03-Jun-2002
 C:Accession: S16221; S14093
 R:Gencic, S.; Schaeffer, H.; von Jagow, G.
 Eur. J. Biochem. 199, 123-131, 1991
 A>Title: Core I protein of bovine ubiquinol-cytochrome-c reductase: an additional me
 re of the protein.
 A:Reference number: S16220; MUID:91293112; PMID:1712295
 A:Accession: S16221
 A:Molecule type: mRNA
 A:Residues: 1-453 <GRN>
 A:Cross-references: GB:X59693; NID:G2399; PID:CA442214.1; PID:G300
 A>Note: Part of this sequence, including the amino end of the mature protein, was co
 R:Cocco, T.; Lorusso, M.; Sardanello, A.M.; Minuto, M.; Ronchi, S.; Tedeschi, G.; Pa
 Eur. J. Biochem. 195, 731-734, 1991
 A>Title: Structural and functional characteristics of polypeptide subunits of the bo
 A:Reference number: S14093; MUID:91153313; PMID:1847870
 A:Accession: S14093
 A:Molecule type: protein
 A:Residues: 17, 16-40, R, 42-46 <COC>
 C:Superfamily: mitochondrial processing peptidase alpha chain
 C:Keywords: hydroxylase; mitochondrial matrix; mitochondrion; oxidative phosphorylatio
 F:1-14/Domain: transmembrane (mitochondrion) #status predicted <TRP>
 F:15-453/Product: ubiquinol-cytochrome-c reductase core protein II #status experimen

Query Match 82.1%; Score 32; DB 1; Length 453;
 Best Local Similarity 66.7%; Pred. No. 11;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
 DB 302 GMANSSSLF 310

RESULT 7
F81313
peptidase (M3 family) Cj1099 [imported] - Campylobacter jejuni (strain NCTC 11168)
C/Species: Campylobacter jejuni
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C/Accession: F81313
R/Parikh, J.; Wren, B.W.; Mungall, K.; Kellley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrett
Nature 403, 665-668, 2000
A/Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyf
A/Reference number: A81250; MUID:20150912; PMID:10688204
A/Accession: F81313
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-573 <PAR>
A/Cross-references: GB:AL139077; GB:AL111168; NID:G6968444; PIDN:CA873354.1; PID:G696853
A/Experimental source: serotype O2, strain NCTC 11168
C/Genetics:
A/Gene: Cj1099
C/Superfamily: oligoendopeptidase F

Query Match 82.1%; Score 32; DB 2; Length 573;
Best Local Similarity 77.8%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GANAXSLF 9
Db 150 GANAFSRLF 158

RESULT 8
S76294
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C/Species: Synechocystis sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
R/Accession: S76294
R/Kaneke, T.; Sato, S.; Kocani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okunura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A/Reference number: S74322; MUID:97061201; PMID:8905231
A/Accession: S76294
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-340 <KAN>
A/Cross-references: EMBL:D64000; GB:AB001339; NID:G1001484; PIDN:BA10146.1; PID:G100151
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C/Genetics:
A/Start codon: GTC
C/Superfamily: hypothetical protein sl10462

Query Match 79.5%; Score 31; DB 2; Length 340;
Best Local Similarity 87.5%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GANAXSL 8
Db 131 GANALSL 138

RESULT 9
T04998
hypothetical protein T16L1.300 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
R/Accession: T04998
R/Beyan, M.; Obermaier, B.; Deutschenhauf, S.; Piravandi, E.; Hohenisel, J.; Mewes, H.W.;
submitted to the Protein Sequence Database, November 1998
A/Reference number: Z15393
A/Accession: T04998
A/Molecule type: DNA

A/Residues: 1-536 <BEV>
A/Cross-references: EMBL:AL031394
A/Experimental source: cultivar Columbia; BAC clone T16L1
C/Genetics:
A/Map position: 4
A/Intons: 54/1, 117/3; 150/3; 415/3
A/Note: T16L1.300

Query Match 79.5%; Score 31; DB 2; Length 536;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GANAXSLF 9
Db 311 GANASSSFF 319

RESULT 10
T06165
multidrug resistance protein 1 homolog - barley
C/Species: Hordeum vulgare (barley)
C/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jul-2000
R/Accession: T06165
R/Davies, T.G.E.; Theodoulou, F.L.; Hallahan, D.L.; Forde, B.G.
Gene 199, 195-202, 1997
A/Title: Cloning and characterization of a novel P-glycoprotein homologue from barley
A/Reference number: Z15500; MUID:98019088; PMID:9358056
A/Accession: T06165
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1232 <DAV>
A/Cross-references: EMBL:Y10099; NID:G2292906; PIDN:CA71179.1; PID:G2292907
A/Experimental source: cv. Maris Mink
C/Superfamily: multidrug resistance protein; ATP-binding cassette homology
F;381-575/Domain: ATP-binding cassette homology <ABC>

Query Match 79.5%; Score 31; DB 2; Length 1232;
Best Local Similarity 66.7%; Pred. No. 55;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GANAXSLF 9
Db 958 GANAVASVF 966

RESULT 11
F83144
conserved hypothetical protein PA4017 [imported] - Pseudomonas aeruginosa (strain PAO)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
R/Accession: F83144
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; L
Lory, S.; Olsen, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: F83144
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-213 <STO>
A/Cross-references: GB:AE004818; GB:AE004091; NID:G9950200; PIDN:AG07404.1; GSPDB:GN
C/Genetics:
A/Gene: PA4017

Query Match 76.9%; Score 30; DB 2; Length 213;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GANAXSLF 9
Db 116 GADKXSLF 124

RESULT 12
H84428
Probable membrane protein [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C/Accession: H84428
R/Link: X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentol, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.;
Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: H84428
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-250 <STO>
A/Cross-references: GB:AE002093; NID:94220472; PIDN:AMD12695.1; GSPDB:GN00139
C/Genetics:
A/Gene: At2g01770
A/Map position: 2

Query Match 76.9%; Score 30; DB 2; Length 250;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GANAXSSL 8
DB 58 GANAXSSI 65

RESULT 13
C89103
Protein C18B10.10 [imported] - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C/Accession: C89103
R/Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A/Reference number: A75000; MUID:99069613; PMID:9851916
A/Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A/Accession: C89103
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-291 <STO>
A/Cross-references: GB:chr_V; PIDN:ACG25811.1; PID:IG1699059; GSPDB:GN00023; CESP:C18B10.
C/Genetics:
A/Gene: C18B10.10
A/Map position: 5

Query Match 76.9%; Score 30; DB 2; Length 291;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 GANAXSLF 9
DB 104 GANSESDLF 112

RESULT 14
C69812
Ferrichrome ABC transporter (permease) homolog yfmd - Bacillus subtilis
C/Species: Bacillus subtilis
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C/Accession: C69812
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berber
A.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, V.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen

lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, P.
Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinc
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portere
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadle, Y.; Sato, T.; Scari
A/Authors: Schleich, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
Aeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiye
T.; Winkler, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida
A/Authors: Yoshikawa, H.F.; Zumestein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: C69812
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-333 <KUN>
A/Cross-references: GB:Z99108; GB:AL009126; NID:92633055; PIDN:CAE12580.1; PID:926330
A/Experimental source: strain 168
C/Genetics:
A/Gene: yfmd
C/Superfamily: ferrichrome ABC transporter

Query Match 76.9%; Score 30; DB 1; Length 333;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 GANAXSLF 9
DB 99 GANAGASLF 107

RESULT 15
T45786
receptor-protein kinase-like protein - Arabidopsis thaliana
A/Alternate names: protein F26013.190
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Mar-2000
C/Accession: T45786
R/Delseny, M.; Berger, C.; Cooke, R.; Grellier, F.; Landie, M.; Mewes, H.W.; Lemcke, I.
submitted to the Protein Sequence Database, December 1999
A/Reference number: 223013
A/Accession: T45786
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-895 <DBL>
A/Cross-references: EMBL:AJ133452
A/Experimental source: cultivar Columbia; BAC clone F26013
C/Genetics:
A/Map position: 3
A/Note: F26013.190
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hon

Query Match 76.9%; Score 30; DB 2; Length 895;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 GANAXSLF 9
DB 122 GANATNSLF 130

Search completed: January 29, 2004, 14:58:20
Job time : 12.1176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:45:44 ; Search time 5.8233 Seconds
(without alignments)
72.678 Million cell updates/sec

Title: US-10-032-950-2
Perfect score: 39
Sequence: 1 GANAXSLF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	84.6	269	1 AKR_CHICK	Q90655 gallus gall
2	33	84.6	406	1 LMP1_MOUSE	P1438 mus musculu
3	33	84.6	407	1 LMP1_BAT	P14562 rattus norv
4	32	82.1	453	1 UCR2_BOVIN	P23002 bos taurus
5	29	74.4	313	1 FIXB_ECO57	O8x27 escherichia
6	29	74.4	313	1 FIXB_ECOLI	P5674 escherichia
7	29	74.4	313	1 FIXB_ECOLI	P31574 escherichia
8	29	74.4	313	1 FIXB_SHIFL	P59675 shigella fl
9	29	74.4	407	1 LMP1_CRIGR	P49129 cricetus
10	29	74.4	870	1 YC60_YEAST	P25623 saccharomyc
11	28	71.8	211	1 YRAR_ECOLI	P45469 escherichia
12	28	71.8	308	1 GPDA_SYNP7	O93522 synechococc
13	28	71.8	439	1 R130_HUMAN	Q9NP22 homo sapien
14	28	71.8	442	1 R130_MOUSE	Q94090 mus musculu
15	28	71.8	453	1 YU48_YEAST	P47125 saccharomyc
16	28	71.8	673	1 SYM_OE18	P59079 oceanobacil
17	28	71.8	926	1 SC24_YEAST	P40482 saccharomyc
18	28	71.8	1200	1 DP3A_YEAST	O9P15 campylobact
19	28	71.8	2483	1 PCX_DROME	P14490 drosophila
20	27	69.2	164	1 YC61_YEAST	P25571 saccharomyc
21	27	69.2	172	1 PHB1_FREDI	P07115 fremyella d
22	27	69.2	187	1 FMK1_ECOLI	P04740 escherichia
23	27	69.2	300	1 PANE_PRRAB	O9V0N0 pyrococcus
24	27	69.2	335	1 E13B_MAIZE	P49237 zea mays (m
25	27	69.2	342	1 GGH_SOYBN	P93164 glycine max
26	27	69.2	384	1 BCAF_HUMAN	P54687 homo sapien
27	27	69.2	389	1 KSS5_ECOLI	O55276 escherichia
28	27	69.2	411	1 LCVB_SYNP7	O55276 synechococc
29	27	69.2	416	1 LMP1_HUMAN	P11279 homo sapien
30	27	69.2	482	1 YBFS_BACSU	O93816 bacillus su
31	27	69.2	631	1 FLGE_HELPJ	O93816 bacillus su
32	27	69.2	718	1 FLGE_HELPJ	O93816 bacillus su
33	27	69.2	718	1 FLGE_HELPJ	P50610 helicobacte

34	27	69.2	820	1 MDL2_YEAST	P33311 saccharomyc
35	27	69.2	1142	1 GIN4_YEAST	O12633 saccharomyc
36	27	69.2	1321	1 AB11_MOUSE	O9G730 mus musculu
37	27	69.2	1492	1 CTRF_SQUAC	P26382 squallus aca
38	27	69.2	1912	1 DICE_HUMAN	O9UPY3 homo sapien
39	27	69.2	4486	1 DYH9_HUMAN	O9NY69 homo sapien
40	26	66.7	110	1 SECG_ECOLI	P33582 escherichia
41	26	66.7	117	1 R8PA_LEPIN	O817K0 leptospira
42	26	66.7	180	1 MP17_LYMTS	P91797 lymanea sta
43	26	66.7	200	1 RK22_MEDSA	P49163 medicago sa
44	26	66.7	225	1 YDUY_ECOLI	P76220 escherichia
45	26	66.7	231	1 P8AF_ECOLI	P12355 spinacia ol

ALIGNMENTS

RESULT 1
AKR_CHICK
ID AKR_CHICK STANDARD: PRT: 269 AA.
AC Q90655;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE homeodomain protein AKR (avian knotted-related protein).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archaeoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=95396587; PubMed=7667102;
RX Ryan A.K., Tejeda M.L., May D.L., Dubocva M., Deeley R.G.;
RT "Isolation and characterization of the chicken homeodomain protein
AKR";
RT
RL Nucleic Acids Res. 23:3252-3259(1995).
CC -1- FUNCTION: BINDS TO THE P' ELEMENT OF THE APOVDLL1 GENE AND
REPRESS ITS TRANSCRIPTION.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE TALE/TGIF HOMEBOX FAMILY.
CC -1- SIMILARITY: Contains 1 homeobox domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL custodian -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC
EMBL: U25353; AAA83567.1; -
PIR: S6433; S58439.
DR InterPro: IPR01356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR ProDom: PD000010; Homeobox; 1.
DR SMART: SMC0389; HOX; 1.
DR PROSITE: PS00027; HOMEBOX 1; FALSE_NEG.
DR PROSITE: PS00721; HOMEBOX 2; 1.
KW Transcription regulation; Repressor; Homeobox; DNA-binding;
KW Nuclear protein.
FT DNA_BIND 35..97 HOMEBOX (TALE-TYPE).
FT FT 35..97
SQ SEQUENCE 269 AA; 29442 MW; 1074355DCC0C2253 CRC64;
Query Match 84.6%; Score 33; DB 1; Length 269;
Best Local Similarity 77.8%; Pred. No. 1.8;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 2
LMP1_MOUSE STANDARD; PRT; 406 AA.
ID LMP1_MOUSE
AC P1438; Q62020;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Lysosome-associated membrane glycoprotein 1 precursor (LAMP-1) (LGP-A)
DE (LGP-120) (CD107A) (P25).
GN LAMP1 OR LAMP-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. PubMed=2142158;
RX MEDLINE=90307738; PubMed=2142158;
RA Granger B.L., Green S.A., Gabel C.A., Howe C.L., Mellman I.,
RA Helenius A.;
RT "Characterization and cloning of lgp110, a lysosomal membrane
RT glycoprotein from mouse and rat cells.";
RL J. Biol. Chem. 265:12036-12043 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RA Hefferman M., Yousefi S., Dennis J.W.;
RL Submitted (FEB-1990) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 25-406 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=88243732; PubMed=3379044;
RA Chen J.W., Cha Y., Yuxsel K.U., Gracy R.W., August J.T.;
RT "Isolation and sequencing of a cDNA clone encoding lysosomal membrane
RT orco-differentiation antigens.";
RL J. Biol. Chem. 263:8754-8758 (1988).
RN [4]
RP DISULFIDE BONDS.
RX MEDLINE=90237040; PubMed=2332434;
RA Afterburn L.M., Bales B.J., August J.T.;
RT "The disulfide structure of mouse lysosome-associated membrane
RT protein 1.";
RL J. Biol. Chem. 265:7419-7423 (1990).
CC -1- FUNCTION: PRESENTS CARBOHYDRATE LIGANDS TO SELECTINS. ALSO
CC IMPLICATED IN TUMOR CELL METASTASIS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Lysosomal.
CC THIS PROTEIN SHUTTLES BETWEEN LYSOSOMES, ENDOSOMES, AND THE PLASMA
CC MEMBRANE.
CC -1- PTM: O- AND N-GLYCOSYLATED, SOME OF THE N-GLYCANS ATTACHED TO
CC LAMP-1 ARE POLYLACTOSAMINOGLYCANS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LAMP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M32015; AAA39428.1; -
CC EMBL: M25244; AAA39869.1; -
CC EMBL: U03861; AAA39411.1; -
CC PIR: A28067; A28067.
CC PIR: A60534; A60534.
CC MGI: MGI:96745; LAMP1.
CC InterPro: IPR002000; LAMP.
CC Pfam: PF01299; LAMP.1.
CC PRINTS: PR00336; LYSASOCTMP.
CC PROSITE: PS00310; LAMP_1; 2.
CC PROSITE: PS00311; LAMP_2; 1.
CC Transmembrane; Glycoprotein; Lysosome; Signal.
FT SIGNAL 1 24
FT CHAIN 25 406 LYSOSOME-ASSOCIATED MEMBRANE GLYCOPROTEIN

FT DOMAIN 25 370
FT TRANSMEM 371 394
FT DOMAIN 395 406
FT DOMAIN 25 188
FT DOMAIN 189 218
FT DOMAIN 219 370
FT DISULFID 35 74
FT DISULFID 149 185
FT DISULFID 222 259
FT DISULFID 327 364
FT CARBOHYD 31 31
FT CARBOHYD 52 52
FT CARBOHYD 58 58
FT CARBOHYD 70 70
FT CARBOHYD 78 78
FT CARBOHYD 97 97
FT CARBOHYD 101 101
FT CARBOHYD 115 115
FT CARBOHYD 159 159
FT CARBOHYD 177 177
FT CARBOHYD 214 214
FT CARBOHYD 219 219
FT CARBOHYD 232 232
FT CARBOHYD 240 240
FT CARBOHYD 252 252
FT CARBOHYD 282 282
FT CARBOHYD 296 296
FT CARBOHYD 311 311
FT CONFLICT 1 10
FT CONFLICT 25 26
FT CONFLICT 385 385
SQ SEQUENCE 406 AA, 43865 MM, C1B3J7354BB9655 CRR64;
Query Match 84.6%; Score 33; DB 1; Length 406;
Best Local Similarity 77.8%; Pred. No. 2.8;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Cy 1 GANXSSLF 9
Db 280 GANXSSLF 288
RESULT 3
LMP1_MOUSE STANDARD; PRT; 407 AA.
ID LMP1_MOUSE
AC P14562; P97620;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Lysosome-associated membrane glycoprotein 1 precursor (LAMP-1) (120
DE kDa lysosomal membrane glycoprotein) (LGP-120) (CD107A).
GN LAMP1 OR LAMP-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. PubMed=3174652;
RX MEDLINE=89017240; PubMed=3174652;
RA Howe C.L., Granger B.L., Hull M., Green S.A., Gabel C.A., Helenius A.,
RA Mellman I.;
RT "Derived protein sequence, oligosaccharides, and membrane insertion
RT of the 120-kDa lysosomal membrane glycoprotein (lgp120):
RT identification of a highly conserved family of lysosomal membrane
RT glycoproteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:7577-7581 (1988).
RN [2]
RP SEQUENCE OF 22-407 FROM N.A.
RX MEDLINE=89153580; PubMed=2920835;
RA Hienno M., Noguchi Y., Sasaki H., Tanaka Y., Furuno K., Kono A.,
RA Sasaki Y., Kato K.;
RT "Isolation and sequencing of a cDNA clone encoding 107 kDa

RT sialoglycoprotein in rat liver lysosomal membranes." ;
 RL FEBS Lett. 244:351-356(1989).
 RN [3]
 RP SEQUENCE OF 283-357 FROM N.A.
 RC STRAIN=Master Kyoto; Tissue=Arctic smooth muscle;
 RL Adams L.A., Meryn I., Schwartz S.M.;
 CC Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PRESENTS CARBOHYDRATE LIGANDS TO SELECTINS. ALSO
 CC IMPLICATED IN TUMOR CELL METASTASIS.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Lysosomal.
 CC THIS PROTEIN SHUTTLES BETWEEN LYOSOMES, ENDOSOMES, AND THE PLASMA
 CC MEMBRANE.
 CC -1- PTM: O- AND N-GLYCOSYLATED. SOME OF THE N-GLYCANS ATTACHED TO
 CC LAMP-1 ARE POLYLACTOSAMINOGLYCANS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE LAMP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.sib.ch).
 CC -----
 DR EMBL: M34959; AAA41525.1; -
 DR EMBL: X14765; CA32873.1; -
 DR EMBL: U75406; AAB19108.1; -
 DR PIR: A30200; A30200.
 DR InterPro: IPR002000; Lamp.
 DR Pfam: PF01299; Lamp; 1.
 DR PRINTS: PR00336; LYSASOCTDMP.
 DR PROSITE: PS00310; LAMP_1; 2.
 DR PROSITE: PS00311; LAMP_2; 1.
 DR Transmembrane; Glycoprotein; Lysosome; Signal.
 KM SIGNAL 1 21
 FT CHAIN 22 407
 FT FT
 FT DOMAIN 22 371
 FT TRANSMEM 372 395
 FT DOMAIN 396 407
 FT DOMAIN 22 189
 FT DOMAIN 190 219
 FT DOMAIN 220 371
 FT DISULFID 36 75
 FT DISULFID 150 186
 FT DISULFID 223 260
 FT DISULFID 328 365
 FT CARBOHYD 32 32
 FT CARBOHYD 59 59
 FT CARBOHYD 71 71
 FT CARBOHYD 79 79
 FT CARBOHYD 102 102
 FT CARBOHYD 116 116
 FT CARBOHYD 125 125
 FT CARBOHYD 145 145
 FT CARBOHYD 160 160
 FT CARBOHYD 178 178
 FT CARBOHYD 215 215
 FT CARBOHYD 220 220
 FT CARBOHYD 223 223
 FT CARBOHYD 241 241
 FT CARBOHYD 271 271
 FT CARBOHYD 283 283
 FT CARBOHYD 297 297
 FT CARBOHYD 312 312
 FT CONFLICT 293 294
 FT CONFLICT 329 329
 FT CONFLICT 356 357
 FT SEQUENCE 407 AA; 43969 MM; 25947490749A768 CR664;
 Query Match 84.6%; Score 33; DB 1; Length 407;
 Best Local Similarity 77.8%; Pred. No. 2.8;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNAXSLSF 9
 Db 261 GNAATSLF 269
 RESULT 4
 ID UC2R_BOVIN STANDARD; PRT; 453 AA.
 AC P23004;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 26-FEB-2003 (Rel. 41, Last annotation update)
 DE Ubiquinol-cytochrome c reductase complex core protein 2, mitochondrial
 DE precursor (EC 1.10.2.2) (Complex III subunit II).
 GN UQCRC2.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=91293112; PubMed=1712295;
 RA Genic S., Schaeffer H., von Jagow G.;
 RT "Core I protein of bovine ubiquinol-cytochrome-c reductase; an
 RT additional member of the mitochondrial protein-processing family.
 RT Cloning of bovine core I and core II cDNAs and primary structure of
 RT the proteins." ;
 RL Eur. J. Biochem. 199;123-131(1991).
 RN [2]
 RP SEQUENCE OF 15-46.
 RX MEDLINE=86298703; PubMed=2841307;
 RA Capaldi R.A., Gonzalez-Halphen D., Zhang Y.-Z., Yanamura W.;
 RT "Complexity and tissue specificity of the mitochondrial respiratory
 RT chain." ;
 RL J. Bioenerg. Biomembr. 20;291-311(1988).
 RN [3]
 RP SEQUENCE OF 15-43.
 RX MEDLINE=89062436; PubMed=2848575;
 RA Gonzalez-Halphen D., Lindorfer W.A., Capaldi R.M.;
 RT "Subunit arrangement in beef heart complex III." ;
 RL Biochemistry 27;7021-7031(1988).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
 RX MEDLINE=97349328; PubMed=9204897;
 RA Xia D., Yu C.A., Kim H., Xia J.Z., Kachurin A.M., Zhang L., Yu L.,
 RA Deisenhofer J.;
 RT "Crystal structure of the cytochrome bcl complex from bovine heart
 RT mitochondria." ;
 RL Science 277;60-66(1997).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
 RX MEDLINE=98316377; PubMed=9651245;
 RA Iwata S., Lee J.W., Okada K., Lee J.K., Iwata M., Raasmussen B.,
 RA Link T.A., Ramaswamy S., Jap B.K.;
 RT "Complete structure of the 11-subunit bovine mitochondrial cytochrome
 RT bcl complex." ;
 RL Science 281;64-71(1998).
 CC -1- FUNCTION: THIS IS A COMPONENT OF THE UBIQUINOL-CYTOCHROME C
 CC REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH
 CC IS PART OF THE MITOCHONDRIAL RESPIRATORY CHAIN. THE CORE PROTEIN
 CC 2 IS REQUIRED FOR THE ASSEMBLY OF THE COMPLEX.
 CC -1- CATALYTIC ACTIVITY: QH(2) + 2 ferricytochrome c = Q + 2
 CC ferrocyclochrome c.
 CC -1- SUBUNIT: BCL COMPLEX CONTAINS 11 SUBUNITS; 3 RESPIRATORY SUBUNITS
 CC (CYTOCHROME B, CYTOCHROME C1, RISKKE PROTEIN), 2 CORE PROTEINS AND
 CC 6 LOW-MOLECULAR WEIGHT PROTEINS.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane; matrix side.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16.
 CC -1- CAUTION: DOES NOT SEEM TO HAVE A PROTEASE ACTIVITY AS IT LACK THE
 CC ZINC-BINDING SITE.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X59693; CAA42214.1; -
 DR PIR; S16221; ZPROCC2.
 DR PDB; 1QCR; 14-OCT-96.
 DR PDB; 1BGY; 23-FEB-99.
 DR PDB; 1BE3; 16-FEB-99.
 DR MEROPS; M6.974; -
 DR InterPro; IPR001431; Peptidase_M16.
 DR Pfam; PF00675; Peptidase_M16; 1.
 DR Pfam; PF05193; Peptidase_M16_C; 1.
 DR PROSITE; PS00143; INSULINASE; 1.
 KM Mitochondrion; Inner membrane; Electron transport; Respiratory chain;
 KM Oxidoreductase; Transit peptide; 3D-structure.
 FT CHAIN 1 453
 FT TRANSIT 1 14
 FT MITOCHONDRION.
 FT UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX
 FT CORE PROTEIN 2.
 FT T -> R.
 FT VARIANT 41 41
 FT STRAND 42 42
 FT TURIN 44 45
 FT STRAND 48 51
 FT STRAND 58 65
 FT HELIX 69 71
 FT TURIN 74 78
 FT HELIX 79 85
 FT TURIN 87 88
 FT STRAND 91 91
 FT HELIX 96 106
 FT TURIN 107 107
 FT STRAND 109 109
 FT STRAND 111 114
 FT STRAND 119 126
 FT HELIX 127 129
 FT HELIX 130 142
 FT STRAND 144 154
 FT HELIX 148 154
 FT TURIN 155 155
 FT HELIX 156 165
 FT TURIN 166 167
 FT HELIX 169 181
 FT TURIN 194 196
 FT HELIX 197 199
 FT HELIX 202 212
 FT HELIX 215 217
 FT STRAND 218 222
 FT HELIX 227 237
 FT TURIN 244 245
 FT STRAND 257 261
 FT STRAND 266 274
 FT TURIN 278 279
 FT HELIX 281 293
 FT STRAND 299 299
 FT TURIN 300 301
 FT HELIX 308 316
 FT STRAND 322 329
 FT STRAND 334 343
 FT TURIN 344 346
 FT HELIX 347 362
 FT TURIN 363 364
 FT TURIN 368 368
 FT HELIX 369 386
 FT HELIX 389 403
 FT TURIN 403 402
 FT HELIX 409 418
 FT HELIX 421 432
 FT TURIN 433 433

FT STRAND 437 442
 FT HELIX 450 453
 SO SEQUENCE 453 AA; 48148 MW; B7C600DA71CD34CF CMC64;
 Query Match 82.1%; Score 32; DB 1; Length 453;
 Best Local Similarity 66.7%; Pred. No. 5.4;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GANAXSLF 9
 Db 302 GSNATSLY 310
 RESULT 5
 ID FIXB_ECO57 STANDARD; PRT; 313 AA.
 AC O8XA27;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE FIXB protein.
 OS FIXB OR 20048 OR ECS0045.
 GN Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 CX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / KMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kubara S., Shiba T., Hariori M., Shinagawa H.;
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 RL Nature 409:529-533(2001).
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / KMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kubara S., Shiba T., Hariori M., Shinagawa H.;
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 RL -1- FUNCTION: Required for anaerobic carnitine reduction. May bring
 RL reductant to cals (By similarity).
 RL -1- PATHWAY: Carnitine metabolism (conversion of carnitine to gamma-
 RL butyrobetaine).
 CC -1- SUBUNIT: Heterodimer of fixa and fixb (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE ETF ALPHA-SUBUNIT / FIXB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE005181; AAG54345.1; -
 DR EMBL; AP002550; BAB33468.1; -
 DR PIR; E85485; E85485.
 DR PIR; E90634; E90634.
 DR HAMAP; MF_010567; 1
 DR InterPro; IPR001308; ETF_alpha.
 DR Pfam; PF00766; ETF_alpha; 1.
 DR PROSITE; PS00696; ETF_ALPHA; 1.
 KM Electron transport; Flavoprotein; FAD; Complete proteome.

FT NP BIND 255 283 FAD (ADP PART) (POTENTIAL).
SQ SEQUENCE 313 AA; 33513 MW; 18ACFEC6F634F6BD CRC64;

Query Match 74.4%; Score 29; DB 1; Length 313;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
DB 270 GANASQITF 278

RESULT 6

FIXB_ECOL6 STANDARD; PRT; 313 AA.

AC P5674;
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE FIXB protein.
GN FIXB OR C0051/C0052.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;

SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liu S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz F.R.,
RA Wodley H.L.T., Domeneberg M.S., Blatterer F.R.,
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).

REVISED.

RA Plunkett G. III,
RL unpublished observations (APR-2003).
CC -1- FUNCTION: Required for anaerobic carnitine reduction. May bring
CC reductant to cals (By similarity).
CC -1- PATHWAY: Carnitine metabolism (conversion of carnitine to gamma-
CC butyrobetaine).
CC -1- SUBUNIT: Heterodimer of fixA and fixB (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ETP ALPHA-SUBUNIT / FIXB FAMILY.
CC -1- CAUTION: When this sequence was assembled, the third base of codon
CC 137 was missed, generating two ORFs instead of one.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: AE016755; -; NOT_ANNOTATED_CDS.

DR HAMAP: MF 01056; -; 1.
DR PROSITE: PS00696; ETP_ALPHA, 1.
KM Electron transport; Flavoprotein; FAD; Complete proteome.
FT NP BIND 255 283 FAD (ADP PART) (POTENTIAL).
SQ SEQUENCE 313 AA; 33512 MW; A9605542004D14D CRC64;

Query Match 74.4%; Score 29; DB 1; Length 313;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
DB 270 GANASQITF 278

RESULT 7
FIXB_ECOLI STANDARD; PRT; 313 AA.

AC P31574;
DT 01-UTL-1993 (Rel. 26, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE FIXB protein.
GN FIXB OR B0042.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;

SEQUENCE FROM N.A.
RC STRAIN=O4:K74;
RX MEDLINE=9606354; PubMed=7473063;
RA Eichler K., Buchet A., Bourgie F., Kleber H.-P.,
RA Mandrand-Berthelot M.-A.,
RT "The fix Escherichia coli region contains four genes related to
RT carnitine metabolism.";
RL J. Basic Microbiol. 35:217-227(1995).

SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=9233497; PubMed=1630901;
RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
RA Isono K., Mizobuchi K., Nakata A.,
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 0-2.4 min region.";
RL Nucleic Acids Res. 20:3305-3308(1992).

SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=9742617; PubMed=9278503;
RA Blatterer F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.,
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).

FUNCTION.
RC STRAIN=K12 / BW25113;
RX PubMed=12061978;
RA "Walt A., Kahn M.L.,
RT reduction in Escherichia coli.";
RL J. Bacteriol. 184:4044-4047(2002).
CC -1- FUNCTION: Required for anaerobic carnitine reduction. May bring
CC reductant to cals.
CC -1- PATHWAY: Carnitine metabolism (conversion of carnitine to gamma-
CC butyrobetaine).
CC -1- SUBUNIT: Heterodimer of fixA and fixB (Probable).
CC -1- SIMILARITY: BELONGS TO THE ETP ALPHA-SUBUNIT / FIXB FAMILY.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to
CC frameshifts.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: X71977; CAA50798.1; -;
DR EMBL: D10463; -; NOT_ANNOTATED_CDS.
DR EMBL: AB000114; AAC73153.1; -;
DR PIR: B64725; B64725.
DR HSSP: P38974; 1EPF.
DR HAMAP: MF 01056; -; 1.
DR EcoGene; EG1563; fixB.

```

DR InterPro: IPR001308; ERF_alpha.
DR Pfam: PF00766; ERF_alpha; 1.
DR PROSITE: PS00696; ERF_ALPHA; 1.
KW Electron transport; Flavoprotein; FAD; Complete proteome.
FT NP BIND 255 283 FAD (ADP PART) (POTENTIAL).
FT CONFLICT 42 42 A -> V (IN REF. 2).
SQ SEQUENCE 313 AA; 3513 MW; 443DBED7F97EDB16 CRC64;

Query Match 74.4%; Score 29; DB 1; Length 313;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Cy 1 GANAXSLF 9
Db 270 GANASQITF 278

RESULT 8
FIXB SHIFL STANDARD; PRT; 313 AA.
AC P59675;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE FIXB protein.
GN FIXB OR SF0039.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OC NCBI_TaxID=623;
[1]
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; Pubmed=12384590;
RA Jin O., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Chang H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RN Nucleic Acids Res. 30:4432-4441(2002).
[2]
RP CONCEPTUAL TRANSLATION.
RA Rivoire C.;
RL Unpublished observations (Apr-2003).
CC -1- FUNCTION: Required for anaerobic carnitine reduction. May bring
CC reductant to carnitine (By similarity).
CC -1- PATHWAY: Carnitine metabolism (conversion of carnitine to gamma-
CC butyrobetaine).
CC -1- SUBUNIT: Heterodimer of fixA and fixB (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ERF ALPHA-SUBUNIT / FIXB FAMILY.
CC -1- CAUTION: This is a conceptual translation; a stop codon in
CC position 264 was translated as Gln to maximize the similarity with
CC orthologs.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE015041; AAN41705.1; ALT_SEQ.
DR HAMAP: MF_01056; -; 1.
DR PROSITE: PS00696; ERF_ALPHA; 1.
KW Electron transport; Flavoprotein; FAD.
FT NP BIND 255 283 FAD (ADP PART) (POTENTIAL).
FT CONFLICT 42 42 A -> V (IN REF. 2).
SQ SEQUENCE 313 AA; 33483 MW; 416CECE6F0994066 CRC64;

Query Match 74.4%; Score 29; DB 1; Length 313;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Cy 1 GANAXSLF 9
Db 270 GANASQITF 278

RESULT 9
LAMP1_CRIGR STANDARD; PRT; 407 AA.
ID LAMP1_CRIGR
AC P49129;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Lysosome-associated membrane glycoprotein 1 precursor (LAMP-1)
DE (lysosomal membrane glycoprotein A) (LGP A).
GN LAMP1 OR LGPA.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OC NCBI_TaxID=10029;
[1]
RN RN
RP SEQUENCE FROM N.A.
RX MEDLINE=97021428; Pubmed=8867788;
RA Uhayakumar S., Granger B.L.;
RT "Cell surface accumulation of overexpressed hamster lysosomal
RT membrane glycoproteins."
RL Cell. Mol. Biol. Res. 41:405-420(1995).
CC -1- FUNCTION: PRESENTS CARBOHYDRATE LIGANDS TO SELECTINS. ALSO
CC IMPLICATED IN TUMOR CELL METASTASIS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Lysosomal.
CC THIS PROTEIN SHUTTLES BETWEEN LYOSOMES, ENDOSOMES, AND THE PLASMA
CC MEMBRANE (BY SIMILARITY).
CC -1- PTM: O- AND N-GLYCOSYLATED; SOME OF THE N-GLYCANS ATTACHED TO
CC LAMP-1 ARE POLYLACTOSAMINOGLYCANS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LAMP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L18986; AAC37682.1; -
DR InterPro: IPR002000; Lamp; 1.
DR Pfam: PF01299; Lamp; 1.
DR PRINTS: PRO0336; LYSASSOCTDMP.
DR PROSITE: PS00310; LAMP_1; 2.
DR PROSITE: PS00311; LAMP_2; 1.
KW Transmembrane; Glycoprotein; Lysosome; Signal.
FT SIGNAL 1 21
FT CHAIN 22 407
FT FT
FT DOMAIN 22 371
FT LAMP1EM 372 395
FT POTENTIAL.
FT DOMAIN 396 407
FT CYTOPLASMIC (POTENTIAL).
FT DOMAIN 22 189
FT FIRST LUMENAL DOMAIN.
FT DOMAIN 190 219
FT HINGE.
FT DOMAIN 220 371
FT SECOND LUMENAL DOMAIN.
FT DISULFID 36 75
FT BY SIMILARITY.
FT DISULFID 150 186
FT BY SIMILARITY.
FT DISULFID 223 260
FT BY SIMILARITY.
FT DISULFID 328 365
FT BY SIMILARITY.
FT CARBOHYD 32 32
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 40 40
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 57 57
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 72 72
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 79 79
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 98 98
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 102 102
FT N-LINKED (GLCNAC...) (POTENTIAL).

```

```

FT CARBOHYD 116 116 N-LINKED (GLCNAC . . .) (POTENTIAL)
FT CARBOHYD 125 125 N-LINKED (GLCNAC . . .) (POTENTIAL)
FT CARBOHYD 145 145 N-LINKED (GLCNAC . . .) (POTENTIAL)
FT CARBOHYD 160 160 N-LINKED (GLCNAC . . .) (POTENTIAL)
FT CARBOHYD 178 178 N-LINKED (GLCNAC . . .) (POTENTIAL)
FT CARBOHYD 215 215 N-LINKED (GLCNAC . . .) (POTENTIAL)
FT CARBOHYD 220 220 N-LINKED (GLCNAC . . .) (POTENTIAL)
FT CARBOHYD 233 233 N-LINKED (GLCNAC . . .) (POTENTIAL)
FT CARBOHYD 241 241 N-LINKED (GLCNAC . . .) (POTENTIAL)
FT CARBOHYD 253 253 N-LINKED (GLCNAC . . .) (POTENTIAL)
FT CARBOHYD 283 283 N-LINKED (GLCNAC . . .) (POTENTIAL)
FT CARBOHYD 297 297 N-LINKED (GLCNAC . . .) (POTENTIAL)
FT CARBOHYD 304 304 N-LINKED (GLCNAC . . .) (POTENTIAL)
FT CARBOHYD 312 312 N-LINKED (GLCNAC . . .) (POTENTIAL)
SQ SEQUENCE 407 AA; 43786 MW; 651002040F86B3D CRC64;

Query Match 74.4%; Score 29; DB 1; Length 407;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy 1 GANAXSLF 9
Db 281 GANXSSLF 289

RESULT 10
YCSO YEAST STANDARD; PRT; 870 AA.
AC P25623; P25622;
DT 01-MAY-1992 (Rel. 22, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 96.1 kDa protein in RIM1-RP614A intergenic region.
GN YCR030C OR YCR30C/YCR29C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxId=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Cedberg H., Hohmann S., Schaff-Gerstenschlaeger I., Huse K.,
RA Zimmermann F.K.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Gromacka R.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: SOME, TO S.POMER SPB4C3.06.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X59720; CAA42322.1; -
DR PIR; S74291; S74291.
DR SGD; S0000626; YCR030C.
KM Hypothetical protein.
SQ SEQUENCE 870 AA; 96125 MW; 3F2CE2F1562E5277 CRC64;

Query Match 74.4%; Score 29; DB 1; Length 870;
Best Local Similarity 75.0%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Cy 1 GANAXSLF 8
Db 755 GANATSLF 762

RESULT 11

```

```

YRAR ECOLI
ID YRAR_ECOLI STANDARD; PRT; 211 AA.
AC P45469;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yrar.
GN YRAR OR B3152.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=562;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
CC -1- SIMILARITY: TO YEAST YER004W.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U18957; AA57955.1; ALT_INIT.
DR EMBL; AE000396; AAC76186.1; ALT_INIT.
DR EcoGene; EG12783; yrar.
KM Hypothetical protein. Complete proteome.
SQ SEQUENCE 211 AA; 23198 MW; 5E3793BADC691FF CRC64;

Query Match 71.8%; Score 28; DB 1; Length 211;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy 1 GANAXSLF 9
Db 113 GANAXSPF 121

RESULT 12
GPDA SYN7
ID GPDA_SYN7 STANDARD; PRT; 308 AA.
AC Q35522;
DT 28-FEB-2003 (Rel. 41, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glyceral-3-phosphate dehydrogenase [NAD(P)-] (EC 1.1.1.94) (NAD(P)-H-
DE dependent)-3-phosphatase dehydrogenase).
GN GSDA OR SEA0004.
OS Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxId=1140;
RN [1]
RP SEQUENCE FROM N.A.
RA Holtman C.K., Socias T., Mohler B.J., Chen Y., Min H., Golden S.S.,
RA Holtman C.K., Socias T., Mohler B.J., Chen Y., Min H., Golden S.S.,
RA Yoderian P.;
RL "Synechococcus elongatus PCC7942 genome sequence, cosmid 7H1.";
RN Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS TO 196 AND 247.
RA Holtman C.K., Socias T., Mohler B.J., Chen Y., Min H., Golden S.S.,
RA Yoderian P., Sandoval P., Gonzalez A., Salinas I.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Sn-glycerol 3-phosphate + NAD(P)(+) =
CC glycerone phosphate + NAD(P)H.

```



```

QY      1 GANAXSL 8
        |||||:|
db      182 GANARSAL 189
        |||||:|
Query Match      71.8%; Score 28; DB 1; Length 308;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

RA Suzuki Y, Nagai K., Sugano S., Ishii S., Kawata-Ho Y., Saito K.,
 RA Yamamoto U., Makematsu A., Nakamura Y., Kojima S., Nagahara K.,
 RA Masuno Y., Ogo T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
 RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.,
 RA RT Human cDNA sequencing project." ;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBD databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusik B., Falder A.A., Rubin G.M., Hong L.,
 RA Scaplehorn M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshilyuki S., Carninci P., Prange C.,
 RA Raha S.S., Locantello N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunatharatne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Halys S.W.,
 RA Viallalon D.K., Muzny D.M., Sodegren E.J., Lu X., Gibbs R.A.,
 RA Bailey J., Heitton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting U., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schnur J., Myers R.W.,
 RA Ratteffeld J.S., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Buchner A., Schein J.E., Jones S.J.M., Marra M.A.;
 RL "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences." ;
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).
 RN [6]
 RP SEQUENCE OF 89-433 FROM N.A.
 RP TISSUE=Bone marrow;
 RA Zhao M., Gu J., Li N., Peng Y., Han Z., Chen Z.;
 RA RT "A novel gene expressed in human bone marrow." ;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBD databases.
 RN [7]
 RP IDENTIFICATION
 RA MEDLINE=21276436; PubMed=21279123;
 RA Koc E.C., Burkhardt W., Blackburn K., Moseley A., Spremull L.L.;
 RT "The small subunit of the mammalian mitochondrial ribosome:
 RT identification of the full complement of ribosomal proteins present." ;
 RL J. Biol. Chem. 276:19363-19374(2001).
 CC -1- SUBUNIT: Component of the mitochondrial ribosome small subunit.
 CC (28S) which comprises a 12S rRNA and about 30 distinct proteins.
 CC -1- STRUCTURAL LOCATION: Mitochondrial.
 CC -1- TISSUE SPECIFICITY: Heart, skeletal muscle, kidney and liver.
 CC Lower expression in placenta and peripheral blood leukocytes.
 CC -1- CAUTION: Ref.6 sequence differs from that shown due to a number of
 CC sequencing errors that include frameshifts.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AF146192; AAF65227.1; -
 DR EMBL; AL355715; CAB90810.1; -
 DR EMBL; AL355716; CAB90811.1; -
 DR EMBL; AL136706; CAB96641.1; -
 DR EMBL; AK074777; BAC11202.1; -
 DR EMBL; BC007735; AAH07735.1; -
 DR EMBL; AF217523; AAF67634.1; ALT_FRAME.
 DR Genew; HGNC:8769; WRPS30.
 DR GO; GO:0006915; P:apoptosis; TAS.
 DR Ribosomal protein; Mitochondrion;
 FT CONFLICT 33 33 S -> C (IN REF. 4 AND 5).
 FT CONFLICT 82 82 I -> V (IN REF. 3).
 FT SEQUENCE 439 AA; 50348 MW; AD0564743E03290C CRC64;

Query Match 71.8%; Score 28; DB 1; Length 439;
 Best Local Similarity 75.0%; Pred. No. 49;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

2 ANAXSLF 9
 326 ANAXSLF 333

Db

RESULT 14
 RT30 MOUSE STANDARD; PRT; 442 AA.
 AC Q2D0G0; O9CYS8; Q9J0Z2;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Mitochondrial 28S ribosomal protein S30 (S30mt) (MRP-S30).
 GN MRPS30.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217651;
 RA Kawai J., Shingawa A., Shidara K., Yoshino M., Itoh M., Ishii Y.,
 RA Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Azawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Giesi C., King B., Koehle H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirral L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Yasuno T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombere P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitlock C., Wilming L.,
 RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Breast;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Parker A.A., Rubin G.M., Hong L.,
 RA Stepien M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uebin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rata S.S., Loquellano N.A., Peters G.D., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huijck S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.C.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scheraga A., Schein J.E., Jones S.D.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE OF 283-442 FROM N.A.
 RC Carim L., Estivill X., Escarceller M., Sunoy L.;

RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
 CC -1 SUBUNIT: Component of the mitochondrial ribosome small subunit
 CC (28S) which comprises a 12S rRNA and about 30 distinct proteins.
 CC -1 SUBCELLULAR LOCATION: Mitochondrial (By similarity).
 CC -1 CAUTION: Ref.1 (BAB28809) sequence differs from that shown due to
 CC frameshifts in positions 75 and 154.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

DR EMBL; AK011476; BAB27644.1; -
 DR EMBL; AK013360; BAB28809.1; ALT_FRAME.
 DR EMBL; BC029784; AAB29784.1; -
 DR EMBL; AL355719; CAB90812.1; -
 DR MGD; MG1:1926237; MRPS30.
 KW Ribosomal protein; Mitochondrion.
 FT CONFLICT 121 121 A -> T (IN REF. 2).
 FT CONFLICT 127 127 P -> SPG (IN REF. 2).
 FT CONFLICT 155 157 RPR -> QPG (IN REF. 2).
 FT CONFLICT 417 417 H -> N (IN REF. 1; BAB28809).
 SQ SEQUENCE 442 AA; 49939 MW; C35A540D3A524BCC CRC64;

Query Match 71.8%; Score 28; DB 1; Length 442;
 Best Local Similarity 75.0%; Pred. No. 49;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

2 ANAXSLF 9
 320 ANAXSLF 327

Db

RESULT 15
 RT30 YEAST STANDARD; PRT; 453 AA.
 AC P4715;
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 18-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 50.8 kDa protein in MIR-STE18 intergenic region.
 GN YOR078W OR Y1840.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c;
 RX MEDLINE=96437976; PubMed=8840504;
 RA Huang M.-E., Mannu V., Chua J.-C., Galibert F.;
 RT "Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open
 RT reading frames and a gene cluster with a counterpart on chromosome
 RT XI";
 RL Yeast 12:865-875(1996).
 RN [2]
 RP -1- SIMILARITY: BELONGS TO THE INDOLAMINE 2,3-DIOXYGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

DR EMBL; Z49578; CAA83606.1; -
 DR EMBL; L47993; AAB39303.1; -
 DR PIR; S57097; S57097.
 DR SGI; S0003839; YOR078W.
 DR GO; GO:0005737; Cytoplasm; IDA.

DR CO; GO:0009435; P:nicotinamide adenine dinucleotide biosynthesis; IMP.
DR InterPro; IPR000898; IDO_fam.
DR Pfam; PF01231; IDO; 1.
DR PROSITE; PS00876; IDO_1; 1.
DR PROSITE; PS00877; IDO_2; 1.
KW Hypothetical protein; Heme.
FT METAL 331 331
FT METAL 375 375
SQ SEQUENCE 453 AA; 50775 MW; A8BF1702F6827EC9 CRC64; IRON (HEME PROXIMAL LIGAND) (POTENTIAL).
IRON (HEME PROXIMAL LIGAND) (POTENTIAL).
Query Match 71.8%; Score 28; DB 1; Length 453;
Best Local Similarity 75.0%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GANAXSL 8
DB 284 GSNAQSL 291

Search completed: January 29, 2004, 14:53:08
Job time : 6.82353 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:46:19 ; Search time 26.8676 Seconds
(without alignments)
86.441 Million cell updates/sec

Title: US-10-032-950-2
Perfect score: 39
Sequence: 1 GANVXSLF 9

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	84.6	189	11 Q922T9	Q922T9 mus musculus
2	33	84.6	406	11 Q8VH34	Q8VH34 mus musculus
3	33	84.6	407	11 Q9DC13	Q9DC13 mus musculus
4	32	82.1	47	16 Q3586	Q3586 strydomoc
5	32	82.1	91	10 Q8W459	Q8W459 arabidopsis
6	32	82.1	573	16 Q9PN05	Q9PN05 campylobacter
7	31	79.5	340	16 Q55632	Q55632 synecocyst
8	31	79.5	415	12 Q81552	Q81552 hepatitis c
9	31	79.5	536	10 Q81897	Q81897 arabidopsis
10	31	79.5	1164	5 Q8WQAO	Q8WQAO caenorhabditis
11	31	79.5	1203	10 Q8S2E4	Q8S2E4 cryza sativ
12	31	79.5	1232	10 Q23998	Q23998 hordeum vul
13	31	79.5	1234	10 Q8GU81	Q8GU81 cryza sativ
14	30	76.9	213	16 Q9HX10	Q9HX10 pseudomonas
15	30	76.9	226	16 Q8PQ96	Q8PQ96 xanthomonas
16	30	76.9	226	16 Q8PDC5	Q8PDC5 xanthomonas

17	30	76.9	250	10 Q9ZU45	Q9ZU45 arabidopsis
18	30	76.9	333	16 Q34933	Q34933 bacillus su
19	30	76.9	334	10 Q9XEW7	Q9XEW7 tritium ae
20	30	76.9	425	16 Q8RCQ9	Q8RCQ9 fusobacteri
21	30	76.9	686	4 Q9HCF9	Q9HCF9 homo sapien
22	30	76.9	737	4 Q8TBB6	Q8TBB6 homo sapien
23	30	76.9	771	11 Q8BXK1	Q8BXK1 mus musculi
24	30	76.9	895	10 Q9SC24	Q9SC24 arabidopsis
25	30	76.9	895	10 Q94C93	Q94C93 arabidopsis
26	30	76.9	1570	5 Q9VF35	Q9VF35 drosophila
27	29	74.4	250	5 Q8W459	Q8W459 arabidopsis
28	29	74.4	265	9 Q37961	Q37961 bacterioph
29	29	74.4	313	16 Q8XA27	Q8XA27 escherichia
30	29	74.4	314	16 Q8X328	Q8X328 ralestoma s
31	29	74.4	328	17 Q8TLL9	Q8TLL9 methanocarc
32	29	74.4	342	16 Q8X3U0	Q8X3U0 escherichia
33	29	74.4	363	5 Q81812	Q81812 toxoplasma
34	29	74.4	388	5 Q9W410	Q9W410 caenorhabdi
35	29	74.4	430	16 Q986D7	Q986D7 rhizobium l
36	29	74.4	457	16 Q8X2X1	Q8X2X1 escherichia
37	29	74.4	457	16 Q8CX24	Q8CX24 escherichia
38	29	74.4	472	5 Q95X15	Q95X15 caenorhabdi
39	29	74.4	476	10 Q9SHG5	Q9SHG5 arabidopsis
40	29	74.4	488	10 Q9FPE8	Q9FPE8 arabidopsis
41	29	74.4	488	16 Q8PQ34	Q8PQ34 xanthomonas
42	29	74.4	495	16 Q8P8S2	Q8P8S2 xanthomonas
43	29	74.4	498	16 Q8P8S7	Q8P8S7 xanthomonas
44	29	74.4	551	10 Q94C51	Q94C51 arabidopsis
45	29	74.4	551	10 Q951I8	Q951I8 arabidopsis

ALIGNMENTS

RESULT 1	Q922T9	PRELIMINARY;	PRT;	189 AA.
AC	Q922T9;			
DT	01-DEC-2001 (TREMREL. 19, Created)			
DT	01-DEC-2001 (TREMREL. 19, Last sequence update)			
DT	01-MAR-2003 (TREMREL. 23, Last annotation update)			
DE	Similar to lysosomal membrane glycoprotein 1 (Fragment).			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Strausberg R.;			
RL	Submitted (AFR-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC006785; AA06785.1; "			
DR	InterPro; IPR002000; Lamp.			
DR	Pfam; PF01299; Lamp; 1.			
DR	PRINTS; PR00316; LYSASOCTDMP.			
DR	PROSITE; PS00310; LAMP_1; 1.			
DR	PROSITE; PS00311; LAMP_2; 1.			
FT	NON TER			
SQ	SSQUENCE 189 AA; 20456 MW; 71F16D69BA4066FA CRC64;			
Query Match	84.6%; Score 33; DB 11; Length 189;			
Best Local Similarity	77.8%; Pred. No. 13;			
Matches	7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	1 GANVXSLF 9			
Db	63 GANVXSLF 71			
RESULT 2	Q8VH34	PRELIMINARY;	PRT;	406 AA.
ID	Q8VH34;			
AC	Q8VH34;			
DT	01-MAR-2002 (TREMREL. 20, Created)			

```

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE LAMP-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Ono K., Han J.
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY069968; AAL58070.1; -.
DR InterPro; IPR02000; LAMP.
DR Pfam; PF01299; LAMP.1.
DR PRINTS; PR00336; LYSASOCTDMP.
DR PROSITE; PS00310; LAMP_1; 2.
DR PROSITE; PS00311; LAMP_2; 1.
SQ SEQUENCE 406 AA; 43879 MW; C1BD373548DFA85 CRC64;

Query Match      84.6%; Score 33; DB 11; Length 406;
Best Local Similarity 77.8%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 GANXSSLP 9
Db 280 GNMASSLF 288

RESULT 3
ID Q9DC13 PRELIMINARY; PRT; 407 AA.
AC Q9DC13.
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Lysosomal membrane glycoprotein 1.
GN LAMP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21085660; PubMed=1121851;
RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Komoto H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okimoto T., Furuno M., Aono H., Baldarelli R., Barth G.,
RA Blake J., Bottelli D., Boujarda N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita W., Gariboldi M.,
RA Guenichon S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima Y., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Segal T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyman-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK004637; BAB23428.1; -.
DR MGI; MGI:96745; LAMP1.
DR InterPro; IPR02000; LAMP.
DR Pfam; PF01299; LAMP.1.
DR PRINTS; PR00336; LYSASOCTDMP.
DR PROSITE; PS00310; LAMP_1; 2.
DR PROSITE; PS00311; LAMP_2; 1.
SQ SEQUENCE 407 AA; 43936 MW; 696D0C79F627DA84 CRC64;

```

```

Query Match      84.6%; Score 33; DB 11; Length 407;
Best Local Similarity 77.8%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 GANXSSLP 9
Db 281 GNMASSLF 289

RESULT 4
ID Q33586 PRELIMINARY; PRT; 47 AA.
AC Q33586.
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE AGRD (AGRD protein).
GN AGRD OR SAV2037 OR SA1842.1 OR SAS066.
OS Staphylococcus aureus (strain Mu50 / ATCC 70699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RA SPECIES=S.aureus; STRAIN=SA502A;
RX MEDLINE=97342847; PubMed=9197262;
RA Ji G., Beavis R., Novick R.P.;
RT "Bacterial interference caused by autoinducing peptide variants.";
RL Science 276:2027-2030(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Itoh T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami K., Hosoyama A.,
RA Mitamura K., Hirakawa H., Kohara S., Goto S., Yabuzaki J., Kato C.,
RA Kanehisa M., Yamashita A., Oshima S., Furuya K., Yoshino C., Shibata T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiratake K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AF001782; AAB63265.1; -.
DR EMBL; AP003364; BAB58199.1; -.
DR EMBL; AP003385; BAB3124.1; -.
RW Complete proteome.
SQ SEQUENCE 47 AA; 5149 MW; 22893764DF54BFA7 CRC64;

Query Match      82.1%; Score 32; DB 16; Length 47;
Best Local Similarity 77.8%; Pred. No. 5.2;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 GANXSSLP 9
Db 24 GNMASSLF 32

RESULT 5
ID Q8W459 PRELIMINARY; PRT; 91 AA.
AC Q8W459.
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical 10.2 kDa protein (Fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eustrodi II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;

```

RN [1]
 RP SEQUENCE FROM N.A.
 RA Souhwick A., Katlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Panh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashiraki Y., Ichida J., Kamiya A., Kawal J.,
 RA Kim C., Lin U., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Theologis A., Davis R.W.,
 RA Ecker J., Theologis A., Davis R.W.,
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY062831; AL32909.1; -
 KW Hypothetical protein.
 FT NON-TER
 SQ SEQUENCE 91 AA; 10166 MW; 23B136FF2CDD3AF1 CRC64;
 QY Query Match 82.1%; Score 32; DB 10; Length 91;
 Best Local Similarity 66.7%; Pred. No. 10;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Db 21 GSNSSSLF 29
 QY 1 GANAXSLF 9
 Db 21 GSNSSSLF 29
 RESULT 6
 Q9PNJ5 PRELIMINARY; PRT; 573 AA.
 ID Q9PNJ5
 AC Q9PNJ5
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Peptidase (M3 family).
 GN C11099.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Campylobacteriaceae; Campylobacter.
 OC NCBI_TaxID=197;
 OX NCBI_TaxID=197;
 RN NCBI_TaxID=197;
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCCT 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parkhill J., Wren B.W., Mungall K., Kellley J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
 RA Jagsis K., Karlyshev A.V., Moule S., Pallen M.V., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Barrall B.G.,
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences."
 RL Nature 403:665-668(2000).
 DR EMBL: AL139077; CAB73354.1; -
 DR InterPro: IPR004438; PeptF.
 DR InterPro: IPR001567; Peptidase_M3.
 DR InterPro: IPR006025; Zn_MTPeptase.
 DR Pfam: PF01432; Peptidase_M3; 1.
 DR TIGRFAMs: TIGR00181; PeptF.1
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KW Complete proteome.
 SQ SEQUENCE 573 AA; 67432 MW; 53A110FFCF0A5A98 CRC64;
 QY Query Match 82.1%; Score 32; DB 16; Length 573;
 Best Local Similarity 77.8%; Pred. No. 73;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Db 1 GANAXSLF 9
 Db 150 GANAFSLF 158
 RESULT 7
 Q55632 PRELIMINARY; PRT; 340 AA.
 ID Q55632
 AC Q55632
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein slr0784.
 GN SLR0784.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN NCBI_TaxID=1148;
 RP SEQUENCE FROM N.A.
 RA STRAIN=PCC 6803;
 RC Tabata S.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96127529; PubMed=8590279;
 RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
 RA Sugita M., Tabata S.,
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
 RT region from map positions 64k to 92k of the genome."
 RL DNA Res. 2:153-166(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirosewa M., Sugita M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsumoto A., Muraki A., Nakazaki N., Natsu K., Okumura S.,
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yanada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions."
 RL DNA Res. 3:109-136(1996).
 DR EMBL: D64000; BAA10146.1; -
 DR InterPro: IPR002549; UPF0118.
 DR Pfam: PF01594; UPF0118; 1.
 KW Hypothetical protein; Complete proteome.
 KM SEQUENCE 340 AA; 37941 MW; 62817BACC97405 CRC64;
 QY Query Match 79.5%; Score 31; DB 16; Length 340;
 Best Local Similarity 87.5%; Pred. No. 71;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 1 GANAXSLF 8
 Db 131 GANALSLF 138
 RESULT 8
 Q81552 PRELIMINARY; PRT; 415 AA.
 ID Q81552
 AC Q81552
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Core, env, and part of E2/NSI (Genome polypeptide) (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN NCBI_TaxID=11103;
 RP SEQUENCE FROM N.A.
 RC STRAIN=NE145;
 RX MEDLINE=9418155; PubMed=8138250;
 RA Shrestha S.M., Tsuda F., Okamoto H., Tokita H., Horikita M.,
 RA Tanaka T., Miyakawa Y., Mayumi M.;
 RT "Hepatitis B virus subtypes and hepatitis C virus genotypes in
 RT patients with chronic liver disease in Nepal."
 RL Hepatology 19:805-809(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NE145;
 RX MEDLINE=94201770; PubMed=8151307;
 RA Tokita H., Shrestha S.M., Okamoto H., Sakamoto M., Horikita M.,

RA Iluka R., Shrestha S., Miyakawa Y., Mayumi M.,
 RT "Hepatitis C virus variants from Nepal with novel genotypes and their
 RT classification into the third major group.";
 RL J. Gen. Virol. 75:931-936(1994).
 CC -1- SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND RNA (BY SIMILARITY).
 CC EMBL; D16618; BAA04040.1; -;
 DR InterPro; IPR002522; HCV_core.
 DR InterPro; IPR002521; HCV_core.
 DR InterPro; IPR002519; HCV_core.
 DR InterPro; IPR002531; HCV_NSI.
 DR Pfam; PF01543; HCV_capsid; 1.
 DR Pfam; PF01542; HCV_core; 1.
 DR Pfam; PF01539; HCV_env; 1.
 DR Pfam; PF01560; HCV_NSI; 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polypeptide; Transmembrane.
 FT NON_TER 415 415
 SQ SEQUENCE 415 AA; 44580 MW; 35FAFF4DF6F79C30 CRC64;
 Query Match 79.5%; Score 31; DB 12; Length 415;
 Best Local Similarity 66.7%; Pred. No. 88;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANXSSLP 9
 DB 389 GSNAGSTLF 397

RESULT 9
 ID 081897 PRELIMINARY; PRT; 536 AA.
 AC 081897;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Beta-xylan endohydrolase -like protein (Beta-xylan endohydrolase-like
 DE protein).
 DE T16L1.300 OR AT4G33810.
 GN Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RX NCB1_TaxID=3702;
 RN SEQUENCE FROM N.A.
 RP Obermaier B., Deutschenbaur S., Piravandi E., Hobeisel J., Jeeze T.,
 RA Heijnen L., Vos P., Mewes H.W., Mayer K.F.X., Schueller C., Bevan M.;
 RL Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.
 RN SEQUENCE FROM N.A.
 RP Obermaier B., Deutschenbaur S., Piravandi E., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN SEQUENCE OF 1-0 FROM N.A.
 RA Vitale D., Liguori R., Argitrou A., De Simone V., Mewes H.W.,
 RA Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN SEQUENCE FROM N.A.
 RP Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AL033394; CAA20594.1; -;
 DR EMBL; AL161584; CAB80098.1; -;
 DR InterPro; IPR001000; Glyco_hydro_10.
 DR Pfam; PF00331; Glyco_hydro_10; 1.
 DR SMART; SM00633; Glyco_10; 1.
 KM Hydrolase.
 SQ SEQUENCE 536 AA; 61259 MW; F46753EC11200225 CRC64;

Query Match 79.5%; Score 31; DB 10; Length 536;
 Best Local Similarity 66.7%; Pred. No. 1,2e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANXSSLP 9
 DB 311 GANXSSFP 319

RESULT 10
 ID 08WQAO PRELIMINARY; PRT; 1161 AA.
 AC 08WQAO;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Y105EBA.24 protein.
 GN Y105EBA.24.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RX NCB1_TaxID=6239;
 RN SEQUENCE FROM N.A.
 RP Sulston J.E.;
 RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
 RN SEQUENCE FROM N.A.
 RP MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology";
 RL Science 282:2012-2018(1998).
 DR EMBL; AL132876; CAD21673.1; -;
 DR WormPep; Y105EBA.24; CE29842.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000219; RhogEF.
 DR Pfam; PF00621; RhogEF; 1.
 DR SMART; SM00233; PH; 1.
 DR PROSITE; PSS0010; DH_2; 1.
 SQ SEQUENCE 1161 AA; 129420 MW; 6E5E474E52E38C1 CRC64;
 Query Match 79.5%; Score 31; DB 5; Length 1161;
 Best Local Similarity 66.7%; Pred. No. 2,6e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANXSSLP 9
 DB 317 GANXSSLP 325

RESULT 11
 ID 08S2E4 PRELIMINARY; PRT; 1203 AA.
 AC 08S2E4;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative P-glycoprotein.
 GN P0022P10.15.
 OS Oryza sativa (Japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriatridae; Oryzae; Oryza.
 OC NCBI_TaxID=39947;
 RX NCB1_TaxID=39947;
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone: p0022p10";
 RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.

DR EMBL: AF003229; BAB89499.1; -
 DR Gramene: Q882E4; -
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001140; ABC_TM_transp.
 DR InterPro: IPR003439; ABC_transporter.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR Pfam: PF00005; ABC_tran; 2.
 DR ProDom: PD000006; ABC_transporter; 2.
 DR SMART: SM00382; AAA; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 DR ATP-binding; Transport; 2.
 KW SEQUENCE 1203 AA; 131507 MW; B7838BDC1B5F19AA CRC64;
 SQ

Query Match 79.5%; Score 31; DB 10; Length 1203;
 Best Local Similarity 66.7%; Pred. No. 2.8e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANXSSLF 9
 DB 929 GANAVASVF 937

RESULT 12
 ID 023998 PRELIMINARY; PRT; 1232 AA.

AC 023998; -
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE P-glycoprotein homologue.
 OS Hordeum vulgare (Barley).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 CC Triticeae; Hordeum.
 OC NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Maria Munk;
 RX MEDLINE=98019088; PubMed=9358056;
 RA Davies T.G.E., Theodouou F.L., Hallahan D.L., Forde B.G.;
 RT "Cloning and characterization of a novel P-glycoprotein homologue from
 RT barley";
 RL Gene 199;195-202 (1997).
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL: Y10099; CA71179.1; -
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001140; ABC_TM_transp.
 DR InterPro: IPR003439; ABC_transporter.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR Pfam: PF00005; ABC_tran; 2.
 DR ProDom: PD000006; ABC_transporter; 2.
 DR SMART: SM00382; AAA; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 DR ATP-binding; Transport; 2.
 KW SEQUENCE 1232 AA; 134377 MW; 17C9604498E42BC5 CRC64;
 SQ

Query Match 79.5%; Score 31; DB 10; Length 1232;
 Best Local Similarity 66.7%; Pred. No. 2.8e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANXSSLF 9
 DB 958 GANAVASVF 966

RESULT 13
 ID 08GUB1 PRELIMINARY; PRT; 1234 AA.

AC 08GUB1; -
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE MDR-like ABC transporter.

GN MDR17.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhardiaceae; Oryzaceae; Oryza.
 OC NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Uasinski M., Ducos E., Martinoia E., Boutry M.;
 RT "The ATP-binding cassette transporters: structure, function and gene
 RT family comparison between rice and Arabidopsis";
 RL Plant Physiol. 0.0-0.0(0).
 DR EMBL: AF535055; CAD59577.1; -
 KW SEQUENCE 1234 AA; 134571 MW; C5F9E9D75D28AC93 CRC64;
 SQ

Query Match 79.5%; Score 31; DB 10; Length 1234;
 Best Local Similarity 66.7%; Pred. No. 2.8e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANXSSLF 9
 DB 960 GANAVASVF 968

RESULT 14
 ID 09HX10 PRELIMINARY; PRT; 213 AA.

AC 09HX10; -
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE Hypothetical protein PA4017.
 GN PA4017.
 OS Pseudomonas aeruginosa.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 CC Pseudomonadaceae; Pseudomonas.
 OC NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Hoover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagron M.,
 RA Gader R.L., Goltz L., Tolentino E., Westbrook-Medman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Polger K.R., Kas A., Ladwig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen";
 RL Nature 406:559-964 (2000).
 DR EMBL: AB004818; AAG07404.1; -
 DR Hypothetical protein; Complete proteome.
 KW SEQUENCE 213 AA; 23164 MW; AD8AFF1F56306F CRC64;
 SQ

Query Match 76.9%; Score 30; DB 16; Length 213;
 Best Local Similarity 66.7%; Pred. No. 74;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANXSSLF 9
 DB 116 GADAKSIF 124

RESULT 15
 ID 08PQ96 PRELIMINARY; PRT; 226 AA.

AC 08PQ96; -
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein XAC0432.
 GN XAC0432.
 OS Xanthomonas axonopodis (pv. citri).


```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SROUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Camnayan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Ciccarelli R.M.B., Coutinho L.L., Curcio-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.U.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Melandis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.P.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezsa R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitejima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL: AE011670; AM35323.1;
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 226 AA; 24311 MW; 1CD4F9E1F4E39DAC CRC64;

Query Match 76.9%; Score 30; DB 16; Length 226;
Best Local Similarity 66.7%; Pred. No. 79;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANAXSEF 9
DB 128 GANAQSRIF 136

```

Search completed: January 29, 2004, 14:56:45
 Job time : 28.6676 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:52:19 ; Search time 11.176 Seconds
(without alignments)
34.252 Million cell updates/sec

Title: US-10-032-950-2

Perfect score: 39

Sequence: 1 GANAXSLF 9

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA:*

1: /cgn2_6/prodata/1/aa/5A.COMB.pep:*

2: /cgn2_6/prodata/1/aa/5B.COMB.pep:*

3: /cgn2_6/prodata/1/aa/6A.COMB.pep:*

4: /cgn2_6/prodata/1/aa/6B.COMB.pep:*

5: /cgn2_6/prodata/1/aa/6C.COMB.pep:*

6: /cgn2_6/prodata/1/aa/6D.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	94.9	9	4	US-09-339-511-2
2	33	84.6	9	4	US-09-339-511-1
3	32	82.1	9	4	US-08-861-476C-3
4	32	82.1	47	4	US-08-861-476C-6
5	31	79.5	56	4	US-09-595-857B-31
6	30	76.9	9	4	US-09-339-511-5
7	30	76.9	9	4	US-09-339-511-6
8	30	76.9	222	4	US-09-328-352-7790
9	30	76.9	258	4	US-09-556-916-30
10	30	76.9	258	4	US-09-556-916-32
11	30	76.9	228	4	US-09-252-991A-21944
12	30	76.9	672	4	US-09-556-916-26
13	30	76.9	672	4	US-09-556-916-28
14	29	74.4	1380	4	US-09-252-991A-25722
15	28	71.8	9	4	US-09-339-511-7
16	28	71.8	226	4	US-09-339-511-7
17	28	71.8	513	4	US-08-908-332-5
18	28	71.8	513	4	US-09-595-857B-30
19	28	71.8	563	2	US-08-714-168-1
20	27	69.2	563	3	US-09-320-721A-1
21	27	69.2	9	4	US-09-339-511-4
22	27	69.2	109	4	US-09-199-637A-89
23	27	69.2	165	4	US-09-519-232-56
24	27	69.2	170	4	US-09-557-921-14
25	27	69.2	283	4	US-09-252-991A-21171
26	27	69.2	329	4	US-08-230-047-7
27	27	69.2	344	4	US-09-252-991A-30554
28	27	69.2	353	1	US-08-073-807A-17

28	27	69.2	380	1	US-08-073-807A-18	Sequence 18, Appl
29	27	69.2	411	1	US-08-399-561-2	Sequence 2, Appl
30	27	69.2	416	1	US-08-073-807A-2	Sequence 2, Appl
31	27	69.2	416	4	US-09-311-764A-26	Sequence 26, Appl
32	27	69.2	482	4	US-09-557-921-2	Sequence 2, Appl
33	26	66.7	20	1	US-07-988-430-59	Sequence 59, Appl
34	26	66.7	20	1	US-08-425-336-57	Sequence 57, Appl
35	26	66.7	20	1	US-08-488-113B-57	Sequence 57, Appl
36	26	66.7	20	1	US-08-477-484B-57	Sequence 57, Appl
37	26	66.7	20	2	US-08-646-360-57	Sequence 57, Appl
38	26	66.7	20	3	US-08-839-765-57	Sequence 57, Appl
39	26	66.7	20	3	US-09-138-389-57	Sequence 57, Appl
40	26	66.7	20	4	US-09-610-838-57	Sequence 57, Appl
41	26	66.7	20	5	PCT-US92-09487-59	Sequence 59, Appl
42	26	66.7	110	4	US-09-462-843A-3	Sequence 3, Appl
43	26	66.7	165	4	US-09-519-232-58	Sequence 58, Appl
44	26	66.7	244	1	US-07-869-933-32	Sequence 32, Appl
45	26	66.7	244	1	US-08-201-879A-3	Sequence 3, Appl

ALIGNMENTS

```

RESULT 1
US-09-339-511-2
Sequence 2, Application US/09339511
Patent No. 6337385
GENERAL INFORMATION:
APPLICANT: Muir, Tom
APPLICANT: Mayville, Patricia
APPLICANT: No. 6337385ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: Ji Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/09/339, 511
CURRENT FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: 60/090, 402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: peptide
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-2
Query Match          94.9%; Score 37; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GANAXSLF 9
Db 1 GANAXSLF 9
RESULT 2
US-09-339-511-1
Sequence 1, Application US/09339511
Patent No. 6337385
GENERAL INFORMATION:
APPLICANT: Muir, Tom
APPLICANT: Mayville, Patricia
APPLICANT: No. 6337385ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: Ji Guangyong

```

```
/ TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
/ FILE OF INVENTION: INTERFERENCE
/ FILE REFERENCE: 600-1-231N
/ CURRENT APPLICATION NUMBER: US/09/339,511
/ PRIOR FILING DATE: 1999-06-24
/ PRIOR APPLICATION NUMBER: 60/090,402
/ PRIOR FILING DATE: 1998-06-24
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO: 1
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic
/ NAME/KEY: VARIANT
/ LOCATION: (5)
/ OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-1
```

```
Query Match      84.6%; Score 33; DB 4; Length 9;
Best Local Similarity 88.9%; Pred. No. 2.5e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GANAXSSLF 9
        |||||
Db      1 GVNAXSSLF 9
```

```
RESULT 3
US-08-861-476C-3
/ Sequence 3, Application US/08861476C
/ Patent No. 6447786
/ GENERAL INFORMATION:
/ APPLICANT: New York University Medical Center
/ TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
/ FILE REFERENCE: 63753/7
/ CURRENT APPLICATION NUMBER: US/08/861,476C
/ CURRENT FILING DATE: 1997-05-22
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO: 3
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
US-08-861-476C-3
```

```
Query Match      82.1%; Score 32; DB 4; Length 9;
Best Local Similarity 77.8%; Pred. No. 2.5e+05;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 GANAXSSLF 9
        |||||
Db      1 GVNAXSSLF 9
```

```
RESULT 4
US-08-861-476C-6
/ Sequence 6, Application US/08861476C
/ Patent No. 6447786
/ GENERAL INFORMATION:
/ APPLICANT: New York University Medical Center
/ TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
/ FILE REFERENCE: 63753/7
/ CURRENT APPLICATION NUMBER: US/08/861,476C
/ CURRENT FILING DATE: 1997-05-22
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO: 6
/ LENGTH: 47
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
```

US-08-861-476C-6

```
Query Match      82.1%; Score 32; DB 4; Length 47;
Best Local Similarity 77.8%; Pred. No. 1.6;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 GANAXSSLF 9
        |||||
Db      24 GVNAXSSLF 32
```

```
RESULT 5
US-09-595-857B-31
/ Sequence 31, Application US/09595857B
/ Patent No. 6495743
/ GENERAL INFORMATION:
/ APPLICANT: Pauli, Robert E.
/ TITLE OF INVENTION: PLANT XYLANASES
/ FILE REFERENCE: 23461201800
/ CURRENT APPLICATION NUMBER: US/09/595,857B
/ CURRENT FILING DATE: 2001-08-28
/ PRIOR APPLICATION NUMBER: 60/141,543
/ PRIOR FILING DATE: 1999-06-28
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 31
/ LENGTH: 536
/ TYPE: PRT
/ ORGANISM: Arabidopsis thaliana
US-09-595-857B-31
```

```
Query Match      79.5%; Score 31; DB 4; Length 536;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 GANAXSSLF 9
        |||||
Db      311 GANAXSSLF 319
```

```
RESULT 6
US-09-339-511-5
/ Sequence 5, Application US/09339511
/ Patent No. 6337385
/ GENERAL INFORMATION:
/ APPLICANT: Muir, Tom
/ APPLICANT: Mayville, Patricia
/ APPLICANT: No. 6337385ick, Richard P.
/ APPLICANT: Beavis, Ronald
/ APPLICANT: Ji, Guangyong
/ TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
/ FILE REFERENCE: 600-1-231N
/ CURRENT APPLICATION NUMBER: US/09/339,511
/ CURRENT FILING DATE: 1999-06-24
/ PRIOR APPLICATION NUMBER: 60/090,402
/ PRIOR FILING DATE: 1998-06-24
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO: 5
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic
/ NAME/KEY: VARIANT
/ LOCATION: (5)
/ OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-5
```

```
Query Match      76.9%; Score 30; DB 4; Length 9;
```

Best Local Similarity 77.8%; Pred. No. 2.5e+05;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
Db 1 GANAXSLF 9

RESULT 7

US-09-339-511-6
; Sequence 6, Application US/09339511
; Patent No. 6337385
; GENERAL INFORMATION:
; APPLICANT: Walt, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. 6337385ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/09/339,511
; CURRENT FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 6
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-6

Query Match 76.9%; Score 30; DB 4; Length 9;
Best Local Similarity 77.8%; Pred. No. 2.5e+05;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
Db 1 GANAXSLF 9

RESULT 8

US-09-328-352-7790
; Sequence 7790, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7790
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7790

Query Match 76.9%; Score 30; DB 4; Length 222;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
Db 120 GANAXSLF 128

RESULT 9

US-09-556-916-30
; Sequence 30, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-30

Query Match 76.9%; Score 30; DB 4; Length 258;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
Db 150 GASALSWF 158

RESULT 10

US-09-556-916-32
; Sequence 32, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-556-916-32

Query Match 76.9%; Score 30; DB 4; Length 258;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
Db 150 GASALSWF 158

RESULT 11

US-09-252-991A-21944
; Sequence 21944, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136

Query Match 71.8%; Score 28; DB 4; Length 9;
Best Local Similarity 77.8%; Pred. No. 2.5e+05;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GANAXSLP 9
| | | | |
| | | | |
Db 1 GVNAXSSAF 9

Search completed: January 29, 2004, 15:03:20
Job time : 11.1176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:47:39 ; Search time 25.5441 Seconds
(without alignments)
73.223 Million cell updates/sec

Title: US-10-032-950-2

Perfect score: 35

Sequence: 1 GANAXSLF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubpaa/US07_PUB_PUB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US07_PUB_PUB.pep.*
6: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
13: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
17: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
18: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	37	94.9	9	US-10-032-950-2
2	33	84.6	9	US-10-032-950-1
3	32	82.1	9	US-10-201-444-3
4	32	82.1	9	US-10-201-444-5
5	30	76.9	9	US-10-032-950-5
6	30	76.9	9	US-10-032-950-6
7	30	76.9	9	US-10-368-687-30
8	30	76.9	9	US-10-368-687-32
9	30	76.9	9	US-10-368-687-26
10	30	76.9	9	US-10-368-687-28
11	30	76.9	9	US-10-108-605-303
12	29	74.4	122	US-09-864-408A-7538
13	29	74.4	235	US-10-083-357-1280
14	29	74.4	491	US-10-094-749-2461
15	29	74.4	550	US-10-156-761-11286

16	29	74.4	1218	US-10-168-428-2	Sequence 2, Appl1
17	28	71.8	9	US-10-032-950-7	Sequence 7, Appl1
18	28	71.8	158	US-10-108-260A-4061	Sequence 4061, Ap
19	28	71.8	159	US-10-407-079-37	Sequence 37, Appl
20	28	71.8	233	US-09-864-761-36239	Sequence 36239, A
21	28	71.8	233	US-09-864-761-42837	Sequence 42837, A
22	28	71.8	305	US-09-925-300-1194	Sequence 1194, Ap
23	28	71.8	389	US-10-291-265-316	Sequence 316, App
24	28	71.8	439	US-10-265-788	Sequence 788, App
25	28	71.8	439	US-10-1295-027-80	Sequence 80, Appl
26	28	71.8	439	US-10-176-847-18	Sequence 18, Appl
27	28	71.8	439	US-10-177-283-355	Sequence 355, App
28	28	71.8	441	US-10-369-493-2265	Sequence 2265, Ap
29	28	71.8	441	US-10-369-493-3165	Sequence 3165, Ap
30	28	71.8	453	US-10-369-493-6581	Sequence 6581, Ap
31	28	71.8	453	US-10-369-493-6582	Sequence 6582, Ap
32	28	71.8	453	US-10-369-493-6583	Sequence 6583, Ap
33	28	71.8	453	US-10-369-493-12645	Sequence 12645, A
34	28	71.8	453	US-10-369-493-12649	Sequence 12649, A
35	28	71.8	465	US-10-369-493-12683	Sequence 12683, A
36	28	71.8	465	US-10-369-493-12683	Sequence 4301, Ap
37	27	69.2	9	US-10-032-950-4	Sequence 4, Appl1
38	27	69.2	61	US-10-029-386-27613	Sequence 27613, A
39	27	69.2	109	US-09-975-719-89	Sequence 89, Appl
40	27	69.2	118	US-10-173-551-14	Sequence 14, Appl
41	27	69.2	170	US-10-328-675A-56	Sequence 56, Appl
42	27	69.2	165	US-10-346-356-14	Sequence 14, Appl
43	27	69.2	286	US-10-156-761-14547	Sequence 14547, A
44	27	69.2	315	US-09-886-055-279	Sequence 279, App
45	27	69.2	315	US-09-804-291-279	Sequence 279, App

ALIGNMENTS

RESULT 1
US-10-032-950-2
Sequence 2, Application US/10032950
Publication No. US20020077453A1
GENERAL INFORMATION:
APPLICANT: Mair, Tom
APPLICANT: Mayville, Patricia
APPLICANT: No. US20020077453A1ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: Ji, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/10/032,950
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-2

Query Match 94.9%; Score 37; DB 13; Length 9;
Best local similarity 100.0%; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GANAXSLF 9
DB 1 GANAXSLF 9

```
RESULT 2
US-10-032-950-1
; Sequence 1, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mullr, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-1

Query Match      84.6% Score 33; DB 13; Length 9;
Best Local Similarity 88.9% Pred. No. 7e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
DB 1 GVNACSLF 9

RESULT 3
US-10-201-444-3
; Sequence 3, Application US/10201444
; Publication No. US20030078378A1
; GENERAL INFORMATION:
; APPLICANT: New York University Medical Center
; TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
; FILE REFERENCE: 63753/7
; CURRENT APPLICATION NUMBER: US/10/201,444
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/08/861,476
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-201-444-3

Query Match      82.1% Score 32; DB 15; Length 9;
Best Local Similarity 77.8% Pred. No. 7e+05;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
DB 1 GVNACSLF 9

RESULT 4
US-10-201-444-6
```

```
; Sequence 6, Application US/10201444
; Publication No. US20030078378A1
; GENERAL INFORMATION:
; APPLICANT: New York University Medical Center
; TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
; FILE REFERENCE: 63753/7
; CURRENT APPLICATION NUMBER: US/10/201,444
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/08/861,476
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-201-444-6

Query Match      82.1% Score 32; DB 15; Length 47;
Best Local Similarity 77.8% Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
DB 24 GVNACSLF 32

RESULT 5
US-10-032-950-5
; Sequence 5, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mullr, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-5

Query Match      76.8% Score 30; DB 13; Length 9;
Best Local Similarity 77.8% Pred. No. 7e+05;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
DB 1 GVNAXSLF 9

RESULT 6
US-10-032-950-6
; Sequence 6, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mullr, Tom
```


APPLICANT: Mayville, Patricia
APPLICANT: No. US2002007453A1ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: Ji, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/10/032,950
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 6
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-6

Query Match 76.9%; Score 30; DB 13; Length 9;
Best Local Similarity 77.8%; Pred. No. 7e+05; 1; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 1;

QY 1 GANAXSLF 9
Db 1 GANAXSLF 9

RESULT 7
US-10-368-687-30
Sequence 30, Application US/10368687
Publication No. US20030162713A1
GENERAL INFORMATION:
APPLICANT: Turner, Alex
APPLICANT: Zambrowicz, Brian
APPLICANT: Nehls, Michael
APPLICANT: Friedrich, Glenn
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. US20030162713A1 Human Transporter Proteins
FILE REFERENCE: 8535-041-999
CURRENT APPLICATION NUMBER: US/10/368,687
CURRENT FILING DATE: 2003-02-14
PRIOR APPLICATION NUMBER: US/09/556,916
PRIOR FILING DATE: 2000-04-21
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 30
LENGTH: 258
TYPE: PRT
ORGANISM: Homo sapiens
US-10-368-687-30

Query Match 76.9%; Score 30; DB 12; Length 258;
Best Local Similarity 66.7%; Pred. No. 88;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
Db 150 GASALSMF 158

RESULT 8
US-10-368-687-32
Sequence 32, Application US/10368687
Publication No. US20030162713A1
GENERAL INFORMATION:
APPLICANT: Turner, Alex

APPLICANT: Zambrowicz, Brian
APPLICANT: Nehls, Michael
APPLICANT: Friedrich, Glenn
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. US20030162713A1 Human Transporter Proteins
FILE REFERENCE: 8535-041-999
CURRENT APPLICATION NUMBER: US/10/368,687
CURRENT FILING DATE: 2003-02-14
PRIOR APPLICATION NUMBER: US/09/556,916
PRIOR FILING DATE: 2000-04-21
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 32
LENGTH: 258
TYPE: PRT
ORGANISM: Homo sapiens
US-10-368-687-32

Query Match 76.9%; Score 30; DB 12; Length 258;
Best Local Similarity 66.7%; Pred. No. 88;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
Db 150 GASALSMF 158

RESULT 9
US-10-368-687-26
Sequence 26, Application US/10368687
Publication No. US20030162713A1
GENERAL INFORMATION:
APPLICANT: Turner, Alex
APPLICANT: Zambrowicz, Brian
APPLICANT: Nehls, Michael
APPLICANT: Friedrich, Glenn
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. US20030162713A1 Human Transporter Proteins
FILE REFERENCE: 8535-041-999
CURRENT APPLICATION NUMBER: US/10/368,687
CURRENT FILING DATE: 2003-02-14
PRIOR APPLICATION NUMBER: US/09/556,916
PRIOR FILING DATE: 2000-04-21
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 26
LENGTH: 672
TYPE: PRT
ORGANISM: Homo sapiens
US-10-368-687-26

Query Match 76.9%; Score 30; DB 12; Length 672;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANAXSLF 9
Db 150 GASALSMF 158

RESULT 10
US-10-368-687-28
Sequence 28, Application US/10368687
Publication No. US20030162713A1
GENERAL INFORMATION:
APPLICANT: Turner, Alex
APPLICANT: Zambrowicz, Brian
APPLICANT: Nehls, Michael
APPLICANT: Friedrich, Glenn
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. US20030162713A1 Human Transporter Proteins
FILE REFERENCE: 8535-041-999
CURRENT APPLICATION NUMBER: US/10/368,687

```
/ CURRENT FILING DATE: 2003-02-14
/ PRIOR APPLICATION NUMBER: US/09/556,916
/ PRIOR FILING DATE: 2000-04-21
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO: 28
/ LENGTH: 672
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-368-687-28
```

```
Query Match          75.9%; Score 30; DB 12; Length 672;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 GANAXSLF 9
Db 150 GASALSWF 158
```

```
RESULT 11
US-10-108-605-303
/ Sequence 303, Application US/10108605
/ Publication No. US20020160934A1
/ GENERAL INFORMATION:
/ APPLICANT: Broadus, Julie
/ APPLICANT: Stem, Lynn
/ APPLICANT: Bachmann, Jane
/ APPLICANT: Kamdar, Kim
/ TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
/ TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
/ FILE REFERENCE: 3113B
/ CURRENT APPLICATION NUMBER: US/10/108,605
/ CURRENT FILING DATE: 2002-03-27
/ PRIOR APPLICATION NUMBER: US 09/761,142
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/176,418
/ PRIOR FILING DATE: 2000-01-14
/ NUMBER OF SEQ ID NOS: 361
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO: 303
/ LENGTH: 1569
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-10-108-605-303
```

```
Query Match          76.9%; Score 30; DB 14; Length 1569;
Best Local Similarity 66.7%; Pred. No. 6.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 GANAXSLF 9
Db 945 GANSSSLW 953
```

```
RESULT 12
US-09-864-408A-7538
/ Sequence 7538, Application US/09864408A
/ Publication No. US20040009474A1
/ GENERAL INFORMATION:
/ APPLICANT: Leach, Martin D.
/ APPLICANT: Shinkels, Richard A.
/ TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Enco
/ FILE REFERENCE: 21402-012
/ CURRENT APPLICATION NUMBER: US/09/864,408A
/ CURRENT FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: 60/206,690
/ PRIOR FILING DATE: 2000-05-24
/ NUMBER OF SEQ ID NOS: 9068
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 7538
/ LENGTH: 122
/ TYPE: PRT
```

```
/ ORGANISM: Homo sapiens
US-09-864-408A-7538
```

```
Query Match          74.4%; Score 29; DB 12; Length 122;
Best Local Similarity 75.0%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 GANAXSSL 8
Db 45 GANATSLAL 52
```

```
RESULT 13
US-10-083-357-1280
/ Sequence 1280, Application US/10083357
/ Publication No. US20030054370A1
/ GENERAL INFORMATION:
/ APPLICANT: Qiandong Zeng et al.
/ TITLE OF INVENTION: Systemic Discovery of New Genes
/ FILE REFERENCE: 032796-090
/ CURRENT APPLICATION NUMBER: US/10/083,357
/ CURRENT FILING DATE: 2002-02-27
/ NUMBER OF SEQ ID NOS: 1346
/ SEQ ID NO: 1280
/ LENGTH: 255
/ TYPE: PRT
/ ORGANISM: Saccharomyces cerevisiae
US-10-083-357-1280
```

```
Query Match          74.4%; Score 29; DB 15; Length 255;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 GANAXSSL 8
Db 140 GANATSLAL 147
```

```
RESULT 14
US-10-094-749-2461
/ Sequence 2461, Application US/10094749
/ Publication No. US20030219741A1
/ GENERAL INFORMATION:
/ APPLICANT: ISOGAI, TAKAO
/ APPLICANT: SUGIYAMA, TOMOYASU
/ APPLICANT: OTSUKI, TETSUJI
/ APPLICANT: WAKAMATSU, AI
/ APPLICANT: SATO, HIROYUKI
/ APPLICANT: ISHII, SHIZUKO
/ APPLICANT: YAMAMOTO, JUN-ICHI
/ APPLICANT: ISONO, YUUKO
/ APPLICANT: HIO, YURI
/ APPLICANT: OTSUKA, KAOBU
/ APPLICANT: NAGAI, KEIICHI
/ APPLICANT: IRIE, RYOTARO
/ APPLICANT: TAMECHIKI, ICHIRO
/ APPLICANT: SEKI, NAOHITO
/ APPLICANT: YOSHIKAWA, TSUTOMU
/ APPLICANT: OTSUKA, MOTOTYUKI
/ APPLICANT: NAGAHARI, KENJI
/ APPLICANT: MASUHO, YASUHIKO
/ TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
/ FILE REFERENCE: 084335/0160
/ CURRENT APPLICATION NUMBER: US/10/094,749
/ CURRENT FILING DATE: 2002-03-12
/ PRIOR APPLICATION NUMBER: 60/350,435
/ PRIOR FILING DATE: 2002-01-24
/ PRIOR APPLICATION NUMBER: JP 2001-328381
/ PRIOR FILING DATE: 2001-09-14
/ NUMBER OF SEQ ID NOS: 3381
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO: 2461
/ LENGTH: 491
```

TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-094-749-2461

Query Match 74.4%; Score 29; DB 12; Length 491;
 Best Local Similarity 55.6%; Pred. No. 2.9e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GANXSSLF 9
 |||:|:|
 Db 2 GANSTSAVF 10

RESULT 15
 US-10-156-761-11286
 ; Sequence 11286, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 11286
 ; LENGTH: 550
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-11286

Query Match 74.4%; Score 29; DB 15; Length 550;
 Best Local Similarity 66.7%; Pred. No. 3.3e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GANXSSLF 9
 |||:|:|
 Db 470 GANSSDLF 478

Search completed: January 29, 2004, 15:01:45
 Job time : 26.544 secs

PA (UYRQ) UNIV ROCKEFELLER.
PA (UYNY) UNIV NEW YORK STATE

XX Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
 CC WPI; 2000-147202/13.
 DR
 XX
 CC New cyclic peptides for treating infections with *Staphylococcus aureus*
 PT
 PT
 XX
 PS Claim 9; Page 26; 37pp; English.

XX This sequence represents a cyclic peptide derived from the *Staphylococcus*
 CC *aureus* AgrD peptide. The invention relates to AgrD derived peptides,
 CC a composition containing a peptide and a carrier, and a method for the
 CC production of the cyclic peptides. The peptide inhibits the agr response,
 CC which is normally associated with the release of virulence factors of
 CC *Staphylococcus aureus*. An AgrD peptide is produced by *S. aureus* that
 CC activates the agr response in strains of a single group, but interferes
 CC with this response in strains of different groups. The peptides and
 CC composition containing them can be used to treat infections by *S. aureus*.

XX
 SQ Sequence 9 AA;
 Query Match 94.6%; Score 35; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. NO. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVAAXSSLF 9
 Db 1 GVAAXSSLF 9

RESULT 2
 ABP53542
 ID ABP53542 standard; peptide; 9 AA.
 XX
 AC ABP53542;
 XX
 DT 13-DEC-2002 (first entry)
 XX
 DE Cyclic peptide SEQ ID NO:3.
 XX
 KM Cyclic peptide; *Staphylococcus aureus*; infection; antibacterial;
 XX agr response inhibitor.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 5 /note= "any amino acid"
 XX
 XX US2002077453-A1.
 XX
 PD 20-JUN-2002.
 XX
 XX 27-DEC-2001; 2001US-0032950.
 XX
 PF 24-JUN-1998; 98US-090402P.
 PR 24-JUN-1999; 99US-0339511.
 XX
 XX (MUIR/) MUIR T W.
 PA (MAYV/) MAYVILLE P.
 PA (NOVI/) NOVICK R P.
 PA (BEAV/) BEAVIS R.
 PA (JIGG/) JI G.
 XX
 PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
 XX
 DR WPI; 2002-681366/73.
 XX
 PT New cyclic peptides, useful for treating *Staphylococcus aureus*
 PT infections
 XX
 PS Claim 9; Page 10; 18pp; English.

XX ABP53540 to ABP53547 represent cyclic peptides (I) from the present
 CC invention. The present invention also describes a method for treating
 CC *Staphylococcus aureus* infection comprising the administration of a
 CC composition comprising (I). (I) has antibacterial activity, and can be
 CC used as an agr gene response inhibitor. The peptides are useful for
 CC treating *S. aureus* infections.

XX
 SQ Sequence 9 AA;
 Query Match 94.6%; Score 35; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. NO. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVAAXSSLF 9
 Db 1 GVAAXSSLF 9

RESULT 3
 AAM50901
 ID AAM50901 standard; peptide; 9 AA.
 XX
 AC AAM50901;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE AgrD-autoinducing cyclic peptide, inhibitor of agr response.
 XX
 KM *Staphylococcus aureus*; AgrD; agr response; inhibitor; antibiotic;
 XX antibacterial; infection; therapy; cyclic.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 5 /note= "any amino acid"
 FT Misc-difference 5 /note= "note linked to residue 9 to form cyclic
 FT peptide"
 FT Misc-difference 9 /note= "note linked to residue 5 to form cyclic
 FT peptide"
 XX
 XX US6337385-B1.
 XX
 PD 08-JAN-2002.
 XX
 PF 24-JUN-1999; 99US-0339511.
 XX
 PR 24-JUN-1998; 98US-090402P.
 XX
 XX (UYRO) UNIV ROCKEFELLER.
 PA (UYNY) UNIV NEW YORK STATE.
 XX
 PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
 XX
 DR WPI; 2002-170774/22.
 XX
 PT Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
 PT interference and for treating *Staphylococcus aureus* infection in a
 PT subject
 XX
 XX Claim 7; Column 19; 18pp; English.

XX The present sequence is that of a novel synthetic cyclic peptide
 CC of the invention that is capable of inhibiting the agr response of
 CC *Staphylococcus aureus*. It is an AgrD-autoinducing peptide, where
 CC AgrD is a secreted agr-encoded peptide and where the agr locus
 CC controls the synthesis of virulence factor and other extracellular
 CC proteins responsible for pathogenicity in *S. aureus*. Preferred
 CC peptides may have the sequence NH₂-X(n)-Z-X(y)-COOH, with a
 CC cyclic bond between the Z residue and COOH other than a thioester

CC bond, where X is an amino acid, an amino acid analogue, a
 CC peptidomimetic or non-amide isostere, Z is a synthetic or a
 CC biosynthetic amino acid, n is 0-10 and y is 1-10. The cyclic bond
 CC is especially a lactam or lactone bond. The thiololactone
 CC structure within native Agri peptides is required for activation
 CC of the agr response. Elimination of the thiol ester component of
 CC the cyclic ring structure can destroy agr response activating
 CC activity while preserving and enhancing inhibitory activity. A
 CC claimed method of preparing a cyclic peptide involves: assembling
 CC a linear peptide chain on to a solid phase resin support;
 CC deprotecting the resulting protected assembled peptide; treating the
 CC deprotected peptide with neutral buffer for a time sufficient to
 CC form the cyclic peptide and cleave the peptide from the support;
 CC and recovering the cyclic peptide. The peptide is useful for
 CC bacterial interference, especially for the treatment of *S. aureus*
 CC infection.

SQ Sequence 9 AA;
 Query Match 94.6%; Score 35; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVAAXSLLF 9
 DB 1 GVAAXSLLF 9

RESULT 4
 ABU05504
 ID ABU05504 standard; Protein; 163 AA.
 XX
 AC ABU05504;
 XX
 DT 08-APR-2003 (first entry)
 XX
 DE M. tuberculosis and M. leprae marker protein #155.
 XX
 KM Mycobacterioses; survival; virulence; protective antigen; vaccine;
 KM mycobacterial disease; tuberculosis; leprosy.
 XX
 OS Mycobacterium tuberculosis.
 OS Mycobacterium leprae.
 XX
 PN W0200274903-A2.
 XX
 PD 26-SEP-2002.
 XX
 PF 22-FEB-2002; 2002MO-IB01973.
 XX
 PR 22-FEB-2001; 2001US-270123P.
 XX
 PA (INSP) INST PASTEUR.
 XX
 PI Cole S;
 XX
 DR WPI; 2002-759885/82.
 XX
 PT Identifying and selecting genes for survival or virulence of
 PT mycobacteria by a comparative genomic analysis of the sequences of
 PT Mycobacterium tuberculosis and M. leprae -
 XX
 PS Claim 17; Page 331; 874pp; English.
 XX
 CC This invention relates to a novel method for identifying essential genes
 CC for survival or virulence of mycobacteria species. The method comprises
 CC aligning the genomic sequence of a first mycobacterium species on a
 CC genomic sequence of a second mycobacterium species and selecting a
 CC polynucleotide sequence that is highly conserved in both genomes with no
 CC counterparts in other bacterial genomic sequences and that corresponds
 CC to an essential gene for the survival or virulence of mycobacterium
 CC species. The method of the invention is useful for detecting M.
 CC tuberculosis or M. leprae infection. The method reduces the number of

CC potential new targets and protective antigens for new drugs and vaccine
 CC compositions to treat and prevent mycobacterial diseases, particularly
 CC tuberculosis and leprosy. The present sequence represents a marker
 CC protein from Mycobacterium tuberculosis and Mycobacterium leprae
 CC identified using the method of the invention.

SQ Sequence 163 AA;
 Query Match 83.8%; Score 31; DB 23; Length 163;
 Best Local Similarity 77.8%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLLF 9
 DB 132 GVAAXSLLF 140

RESULT 5
 ABU05887
 ID ABU05887 standard; Protein; 167 AA.
 XX
 AC ABU05887;
 XX
 DT 08-APR-2003 (first entry)
 XX
 DE M. tuberculosis and M. leprae marker protein #538.
 XX
 KM Mycobacterioses; survival; virulence; protective antigen; vaccine;
 KM mycobacterial disease; tuberculosis; leprosy.
 XX
 OS Mycobacterium tuberculosis.
 OS Mycobacterium leprae.
 XX
 PN W0200274903-A2.
 XX
 PD 26-SEP-2002.
 XX
 PF 22-FEB-2002; 2002MO-IB01973.
 XX
 PR 22-FEB-2001; 2001US-270123P.
 XX
 PA (INSP) INST PASTEUR.
 XX
 PI Cole S;
 XX
 DR WPI; 2002-759885/82.
 XX
 PT Identifying and selecting genes for survival or virulence of
 PT mycobacteria by a comparative genomic analysis of the sequences of
 PT Mycobacterium tuberculosis and M. leprae -
 XX
 PS Claim 17; Page 748; 874pp; English.
 XX
 CC This invention relates to a novel method for identifying essential genes
 CC for survival or virulence of mycobacteria species. The method comprises
 CC aligning the genomic sequence of a first mycobacterium species on a
 CC genomic sequence of a second mycobacterium species and selecting a
 CC polynucleotide sequence that is highly conserved in both genomes with no
 CC counterparts in other bacterial genomic sequences and that corresponds
 CC to an essential gene for the survival or virulence of mycobacterium
 CC species. The method of the invention is useful for detecting M.
 CC tuberculosis or M. leprae infection. The method reduces the number of
 CC potential new targets and protective antigens for new drugs and vaccine
 CC compositions to treat and prevent mycobacterial diseases, particularly
 CC tuberculosis and leprosy. The present sequence represents a marker
 CC protein from Mycobacterium tuberculosis and Mycobacterium leprae
 CC identified using the method of the invention.

SQ Sequence 167 AA;
 Query Match 83.8%; Score 31; DB 23; Length 167;
 Best Local Similarity 77.8%; Pred. No. 25;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSSLF 9
 |||||
 DB 136 GVAAXSSAF 144

RESULT 6
 ABP57498
 ID ABP57498 standard; Protein; 167 AA.

XX ABP57498;

XX 28-APR-2003 (first entry)

XX Mycobacterium tuberculosis protein SEQ ID NO.125.

XX Mycobacterium tuberculosis; infection; antibacterial; tuberculostratic;

XX immunostimulant; vaccine; gene therapy; mycobacterial infection.

XX Mycobacterium tuberculosis.

XX WO200300721-A2.

XX 03-JAN-2003.

XX 21-JUN-2002; 2002MO-GB02845.

XX 22-JUN-2001; 2001GB-0015365.

XX 07-SEP-2001; 2001GB-0021780.

XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.

XX James BW, Bacon J, Marsh P;

XX WPI; 2003-201403/19.

XX N-PSDB; AB271124.

XX New mycobacterial peptide, its fragment, variant or derivative, useful

XX as vaccine for treating or preventing mycobacterial infections, and as

XX diagnostic reagents for identifying such infections

XX Claim 2; Page 230; 246pp; English.

XX ABP57436 to ABP57504 represent mycobacterial amino acid sequences (I)

XX encoded by AB271062 to AB271130 (II), which are isolated from

XX Mycobacterium tuberculosis. (I) are encoded by genes (II) whose

XX expression is induced or up-regulated during culture of a mycobacterium

XX under conditions defined by a dissolved oxygen tension of at least 10%

XX air saturation measured at 37 plus degrees Celsius, when compared with a

XX dissolved oxygen tension of at least 40% air saturation measured at 37

XX plus degrees Celsius. (I) and (II) have antibacterial, tuberculostatic

XX and immunostimulant activities, and can be used in vaccines and gene

XX therapy. (I) and (II) can be used for the manufacture of a medicament

XX for treating or preventing a mycobacterial infection. They can also be

XX used for the manufacture of a diagnostic reagent for identifying a

XX mycobacterial infection.

XX Sequence 167 AA;

XX Query Match 83.8%; Score 31; DB 24; Length 167;

XX Best Local Similarity 77.8%; Pred. No. 25;

XX Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSSLF 9

XX 136 GVAAXSSAF 144

RESULT 7

AAV67859

ID AAV67859 standard; peptide; 9 AA.

XX AAV67859;

XX 25-APR-2000 (first entry)

XX Staphylococcus aureus AgdII derived peptide sequence.

XX Staphylococcus aureus infection; treatment; Agd; agr response;

XX virulence factor.

XX Staphylococcus aureus.

XX WO9967286-A2.

XX 29-DEC-1999.

XX 24-JUN-1999; 99MO-US14562.

XX 24-JUN-1998; 98US-0103438.

XX (UYRQ) UNIV ROCKEFELLER.

XX (UYNY) UNIV NEW YORK STATE.

XX Muir TW, Mayville P, Novick RP, Ji G, Beavis R;

XX WPI; 2000-147202/13.

XX New cyclic peptides for treating infections with Staphylococcus aureus

XX Examples; Page 22; 37pp; English.

XX This sequence represents the Staphylococcus aureus AgdII derived

XX peptide. The invention relates to Agd derived cyclic peptides, a

XX composition containing a peptide and a carrier, and a method for the

XX production of the cyclic peptides. The peptide inhibits the agr response,

XX which is normally associated with the release of virulence factors of

XX Staphylococcus aureus. An Agd peptide is produced by S. aureus that

XX activates the agr response in strains of a single group, but interferes

XX with this response in strains of different groups. The peptides and

XX composition containing them can be used to treat infections by S. aureus.

XX Sequence 9 AA;

XX Query Match 81.1%; Score 30; DB 21; Length 9;

XX Best Local Similarity 77.8%; Pred. No. 9.3e+05;

XX Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSSLF 9

XX 136 GVAAXSSAF 144

RESULT 8

AAV67860

ID AAV67860 standard; peptide; 9 AA.

XX AAV67860;

XX 25-APR-2000 (first entry)

XX Staphylococcus aureus AgdII derived peptide sequence.

XX Staphylococcus aureus infection; Agd; agr response; treatment;

XX virulence factor.

XX Staphylococcus aureus.

XX WO9967286-A2.

XX 29-DEC-1999.

XX 24-JUN-1999; 99MO-US14562.

XX 24-JUN-1998; 98US-0103438.

```

XX XX (UYRO ) UNIV ROCKEFELLER.
PA PA (UYNV ) UNIV NEW YORK STATE.
XX XX
XX XX Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
XX XX WPI; 2000-147202/13.
XX XX
PT PT New cyclic peptides for treating infections with Staphylococcus aureus
PT PT
XX XX
XX PS Examples; Page 22; 37pp; English.
CC CC This sequence represents the Staphylococcus aureus AgrDII derived
CC peptide. The invention relates to Agrd derived cyclic peptides, a
CC composition containing a peptide and a carrier, and a method for the
CC production of the cyclic peptides. The peptide inhibits the agr response,
CC which is normally associated with the release of virulence factors of
CC Staphylococcus aureus. An Agrd peptide is produced by S. aureus that
CC activates the agr response in strains of a single group, but interferes
CC with this response in strains of different groups. The peptides and
CC composition containing them can be used to treat infections by S. aureus.
XX XX
SQ Sequence 9 AA;
Query Match 81.1%; Score 30; DB 21; Length 9;
Best Local Similarity 77.8%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0
OY 1 GVAAKSSLF 9
DB 1 GVNAASSSLF 9
RESULT 9
AAM50907 ID AAM50907 standard; Peptide; 9 AA.
XX AC AAM50907;
XX DT 08-MAY-2002 (first entry)
XX DE Protected peptide used in cyclic peptide production.
XX KM Staphylococcus aureus; Agrd; agr response: inhibitor; antibiotic;
XX KM antibacterial; infection; therapy; cyclic.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FH FT Modified-site 1 /note= "Z-Gly"
FT FT Modified-site 5 /note= "Ser(Bbu)"
FT FT Modified-site 6 /note= "Ser(BzI)"
FT FT Modified-site 7 /note= "Ser(BzI)"
FT FT Misc-difference 5 /note= "Ser(BzI)"
FT FT Misc-difference 9 /note= "note linked to residue 9 to form cyclic peptide"
FT FT Misc-difference 9 /note= "note linked to residue 5 to form cyclic peptide"
XX PN US6337385-B1.
XX PD 08-JAN-2002.
XX PF 24-JUN-1999; 99US-0339511.
XX PR 24-JUN-1998; 98US-090402P.
XX XX

```

[illegible]


```

PR 24-JUN-1998; 98US-090402P.
XX
XX (UYRO ) UNIV ROCKEFELLER.
PA (UYNY ) UNIV NEW YORK STATE.
XX
XX
XX Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX
XX WPI; 2002-170774/22.
XX
XX
XX Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
PT interference and for treating Staphylococcus aureus infection in a
PT subject
XX
XX Example 1; Column 9; 18pp; English.
XX
XX The present sequence is that of a novel synthetic AgrD2 linear
CC thioester peptide. The peptide is derived from the cyclic AgrD2
CC peptide of Staphylococcus aureus group II. AgrD2 is a secreted
CC agr-encoded peptide, where the agr locus controls the synthesis of
CC virulence factor and other extracellular proteins responsible for
CC pathogenicity in S. aureus. The biological activity of the
CC synthetic peptide was assayed using cultured S. aureus strains
CC containing a beta-lactamase reporter gene fused to the agr3
CC promoter. This allowed activation or inhibition of the agr
CC response to be monitored spectrophotometrically. Unlike an AgrD2
CC thioester cyclic peptide (see AAM51001), the present peptide
CC was unable to either activate or inhibit the agr response, even
CC when added to cultured cells at 1M concentrations. The invention
CC provides claimed cyclic peptides (see AAM50899-906 and AAM50999)
CC and methods for preparing them. The cyclic peptides are useful for
CC bacterial interference, especially for the treatment of S. aureus
CC infection.
XX
XX SQ Sequence 9 AA;
XX
XX Query Match 81.1%; Score 30; DB 23; Length 9;
XX Best Local Similarity 77.8%; Pred. No. 9.3e+05;
XX Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 GVAXSSLF 9
XX |||||
DB 1 GVNAASSLF 9
XX |||||
XX
XX RESULT 11
XX AAM51003
XX ID AAM51003 standard; Peptide; 9 AA.
XX
XX AC AAM51003;
XX
XX DT 08-MAY-2002 (first entry)
XX
XX DE AgrD2 linear free acid peptide.
XX
XX DE AgrD2; agr response; inhibitor; antibiotic; antibacterial;
XX infection; therapy.
XX
XX KM Staphylococcus aureus.
XX
XX OS Synthetic.
XX
XX OS US6337385-B1.
XX
XX PN 08-JAN-2002.
XX
XX PD 24-JUN-1999; 99US-0339511.
XX
XX PF 24-JUN-1999; 98US-090402P.
XX
XX PR 24-JUN-1998; 98US-090402P.
XX
XX PA (UYRO ) UNIV ROCKEFELLER.
XX PA (UYNY ) UNIV NEW YORK STATE.
XX
XX PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX

```

```

DR WPI; 2002-170774/22.
XX
XX Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
PT interference and for treating Staphylococcus aureus infection in a
PT subject
XX
XX Example 1; Column 9; 18pp; English.
XX
XX The present sequence is that of a novel synthetic AgrD2 linear
CC free acid peptide. The peptide is derived from the cyclic AgrD2
CC peptide of Staphylococcus aureus group II. AgrD2 is a secreted
CC agr-encoded peptide, where the agr locus controls the synthesis of
CC virulence factor and other extracellular proteins responsible for
CC pathogenicity in S. aureus. The biological activity of the
CC synthetic peptide was assayed using cultured S. aureus strains
CC containing a beta-lactamase reporter gene fused to the agr3
CC promoter. This allowed activation or inhibition of the agr
CC response to be monitored spectrophotometrically. Unlike an AgrD2
CC thioester cyclic peptide (see AAM51001), the present peptide
CC was unable to either activate or inhibit the agr response, even
CC when added to cultured cells at 1M concentrations. The invention
CC provides claimed cyclic peptides (see AAM50899-906 and AAM50999)
CC and methods for preparing them. The cyclic peptides are useful for
CC bacterial interference, especially for the treatment of S. aureus
CC infection.
XX
XX SQ Sequence 9 AA;
XX
XX Query Match 81.1%; Score 30; DB 23; Length 9;
XX Best Local Similarity 77.8%; Pred. No. 9.3e+05;
XX Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 GVAXSSLF 9
XX |||||
DB 1 GVNAASSLF 9
XX |||||
XX
XX RESULT 12
XX AAM51004
XX ID AAM51004 standard; Peptide; 9 AA.
XX
XX AC AAM51004;
XX
XX DT 08-MAY-2002 (first entry)
XX
XX DE AgrD2 lactone cyclic peptide.
XX
XX DE AgrD2; agr response; inhibitor; antibiotic; antibacterial;
XX infection; therapy; lactone; cyclic.
XX
XX KM Staphylococcus aureus.
XX
XX OS Synthetic.
XX
XX OS US6337385-B1.
XX
XX PN 08-JAN-2002.
XX
XX PD 24-JUN-1999; 99US-0339511.
XX
XX PF 24-JUN-1999; 98US-090402P.
XX
XX PR 24-JUN-1998; 98US-090402P.
XX
XX PA (UYRO ) UNIV ROCKEFELLER.
XX PA (UYNY ) UNIV NEW YORK STATE.
XX
XX PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX

```

XX WPI; 2002-170774/22.
 DR Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
 XX interference and for treating *Staphylococcus aureus* infection in a
 PT subject -
 XX
 XX Example 1; Column 9; 18pp; English.
 XX The present sequence is that of a novel synthetic AgrD2
 CC lactone cyclic peptide in which residue 5 of the peptide is
 CC linked to residue 9 via a lactone bond. The peptide is derived
 CC from an AgrD2 peptide of *Staphylococcus aureus* group II. AgrD2 is
 CC a secreted agr-encoded peptide, where the agr locus controls the
 CC synthesis of virulence factor and other extracellular proteins
 CC responsible for pathogenicity in *S. aureus*. The biological
 CC activity of the synthetic peptide was assayed using cultured *S.*
 CC aureus strains containing a beta-lactamase reporter gene fused to
 CC the agrP3 promoter. This allowed activation or inhibition of the
 CC agr response to be monitored spectrophotochemically. The lactone
 CC AgrD2 peptide inhibited the agr response of group I *S. aureus*
 CC strains without activating the agr response in group I, II or III
 CC strains. The invention provides claimed cyclic peptides (see
 CC AAM5089-906 and AAM50993) and methods for preparing them,
 CC especially peptides where the cyclic bond is a lactam or lactone
 CC bond. The cyclic peptides are useful for bacterial interference,
 CC especially for the treatment of *S. aureus* infection.
 XX
 SQ Sequence 9 AA;
 Query Match 81.1%; Score 30; DB 23; Length 9;
 Best Local Similarity 77.8%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 GVAAXSSLF 9
 DB 1 GVNASSSLF 9
 RESULT 13
 ABB07160
 ID ABB07160 standard; peptide; 9 AA.
 XX
 AC ABB07160;
 XX
 DT 13-MAR-2002 (first entry)
 XX
 DE Peptide-mediated quorum sensing inhibitor peptide cyclo-XII.
 XX
 KW Autoinducer-2; AI-2; antibiotic; antibacterial; dermatological;
 KM vulnerrary; pheromone; agr system; accessory gene regulator; cyclic.
 OS Synthetic.
 XX
 PN WO200185664-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 10-MAY-2001; 2001WO-US15221.
 XX
 PR 10-MAY-2000; 2000US-203000P.
 PR 07-DEC-2000; 2000US-254398P.
 XX
 PA (UYPK-) UNIV PRINCETON.
 PA (UQOR-) QUOREX PHARM INC.
 PA (UYTE-) UNIV TECHNOLOGIES INT INC.
 XX
 PT Bassler BL, Dammel CS, Schauder S, Shokat K, Stein J, Surette MG;
 XX WPI; 2002-075235/10.
 XX Use of autoinducer-2 agonists or antagonists for regulating activity of
 PT autoinducer-2 receptor, regulating bacterial growth and pathogenesis.

PT also antibiotic compositions -
 XX
 XX Disclosure; Page 33; 134pp; English.
 XX
 CC The invention relates to the use of autoinducer-2 (AI-2) agonists or
 CC antagonists for regulating activity of autoinducer-2 receptor, regulating
 CC bacterial growth and pathogenesis. Synergistic antibiotic compositions
 CC comprising inhibitors of the quorum-sensing pathway of a microorganism
 CC are also provided. Methods using such AI-2 analogues are useful for
 CC treating pathogen-associated disease states. The compounds and antibiotic
 CC compositions can be used to inhibit bacterial cell growth and/or biofilm
 CC formation on a medical device, particularly for promoting growth of skin
 CC graft replacements used in the treatment of burns and ulcers. They may
 CC also be used to aid wound repair, and to inhibit bacterial cell growth
 CC and biofilm formation in or on products or devices used for personal
 CC hygiene. The present sequence represents a inhibitor of peptide-mediated
 CC quorum sensing.
 XX
 SQ Sequence 9 AA;
 Query Match 81.1%; Score 30; DB 23; Length 9;
 Best Local Similarity 77.8%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 GVAAXSSLF 9
 DB 1 GVNASSSLF 9
 RESULT 14
 ABG15689
 ID ABG15689 standard; Protein; 28 AA.
 XX
 AC ABG15689;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #15680.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS79876.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 46048; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG0377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 28 AA;

Query Match

Best Local Similarity 81.1%; Score 30; DB 22; Length 28;
 Best Local Similarity 77.8%; Pred. No. 6.5;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9

Db 18 GVAAGMSLF 26

RESULT 15

AAV33146

ID AAV33146 standard; Protein; 543 AA.

XX AAV33146;

DT 16-NOV-1999 (first entry)

DE Rabbit liver carboxylesterase protein fragment.

XX Carboxylesterase; rabbit; liver; tumor cell; chemotherapy; produg;

KM disease-specific responsive promoter; CRT-11; APC; resection;

XX recurrence; inhibition; bone marrow cell.

OS Oryctolagus cuniculus.

XX MO9942533-A1.

XX 26-AUG-1999.

PF 12-FEB-1999; 99WO-US03171.

XX 19-FEB-1998; 98US-0075258.

XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.

PI Danks MK, Potter PM, Houghton PJ;

XX WPI; 1999-540311/45.

XX New carboxylesterase polypeptide for treatment of tumors -

PS Claim 3; Page 63-65; 70pp; English.

CC This invention describes a novel rabbit carboxylesterase polypeptide (I)
 CC capable of metabolizing a chemotherapeutic produg and its inactive
 CC metabolites into an active drug. The invention also describes a
 CC composition comprising (I) and a disease-specific responsive promoter.
 CC This composition is useful for sensitizing tumor cells to a
 CC chemotherapeutic produg (preferably CRT-11 or APC) by transfecting into
 CC tumor cells, and contacting the sensitized cells with a chemotherapeutic
 CC produg to inhibit growth of the tumor cells. The composition can also be
 CC administered to the site of tumor resection to inhibit tumor recurrence,
 CC and be administered to bone marrow cells to remove tumor cells. The
 CC products of the invention are useful for identifying drugs that are
 CC inactivated by a carboxylesterase enzyme, and are also useful for

CC identifying compounds containing a COOC ester linkage that are activated
 CC by a carboxylesterase enzyme. This sequence represents a rabbit liver
 CC carboxylesterase fragment which is described in the method of the
 CC invention.

XX SQ Sequence 543 AA;

Query Match

Best Local Similarity 81.1%; Score 30; DB 20; Length 543;
 Best Local Similarity 77.8%; Pred. No. 1.5e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9

Db 248 GVALLSLF 256

Search completed: January 29, 2004, 14:52:11
 Job time : 36.0735 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OW protein - protein search, using sw model

Run on: January 29, 2004, 14:46:49 ; Search time 11.1176 Seconds
(without alignments)
77.851 Million cell updates/sec

Title: US-10-032-950-3

Perfect score: 37

Sequence: 1 GVAXSLF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 segs, 9616862 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR 76:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	89.2	1253	2	T18528
2	31	83.8	163	2	A87066
3	31	83.8	167	2	C70866
4	31	83.8	169	2	T45243
5	30	81.1	244	2	T43566
6	30	81.1	244	2	A40049
7	30	81.1	267	2	B83602
8	30	81.1	305	2	AC2730
9	30	81.1	305	2	D97511
10	30	81.1	539	2	A29923
11	29	78.4	367	2	E64399
12	29	78.4	368	2	F85056
13	29	78.4	1668	1	C69224
14	29	75.7	42	2	I51291
15	28	75.7	47	2	C89895
16	28	75.7	282	2	T15556
17	28	75.7	330	2	D82756
18	28	75.7	331	2	S33423
19	28	75.7	338	1	Q8SHUA
20	28	75.7	338	2	T43918
21	28	75.7	339	2	AH0359
22	28	75.7	348	2	T12290
23	28	75.7	348	2	T12291
24	28	75.7	363	1	ADRAA
25	28	75.7	364	1	ADHUA
26	28	75.7	364	1	ADHUA
27	28	75.7	364	1	ADHUA
28	28	75.7	382	1	S33630
29	28	75.7	386	2	S21974

30	28	75.7	386	2	S17917	ADP,ATP carrier pr
31	28	75.7	387	2	S16568	ADP,ATP carrier pr
32	28	75.7	387	2	S14876	ADP,ATP carrier pr
33	28	75.7	396	2	AH2859	MPS permease (impo
34	28	75.7	396	2	P97363	probable transport
35	28	75.7	442	2	AF2539	manganese transport
36	28	75.7	523	2	T07834	hydroxymethylpyrim
37	28	75.7	525	2	E86363	probable thiamin b
38	28	75.7	549	2	T23422	hypothetical prote
39	28	75.7	612	2	T23420	hypothetical prote
40	28	75.7	639	2	T23424	hypothetical prote
41	28	75.7	981	2	B88794	protein K07F5.12a
42	27	73.0	182	2	AB1251	hypothetical prote
43	27	73.0	184	2	T46711	hypothetical prote
44	27	73.0	188	2	B38116	hypothetical prote
45	27	73.0	188	2	AC2563	transposase alt850

ALIGNMENTS

RESULT 1

T18528
Probable pyruvate (flavodoxin) dehydrogenase (EC 1.2.99.-) - Giardia intestinalis
N/Alternate names: pyruvate:ferredoxin oxidoreductase
C/Species: Giardia intestinalis
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 15-Sep-2000
C/Accession: T18528
R/Townson, S.M.; Hanson, G.R.; Upcroft, J.A.; Upcroft, P.
Eur. J. Biochem. 220, 439-446, 1994
A/Title: A purified ferredoxin from Giardia duodenalis.
A/Reference number: S42359; MUID:94170792; PMID:8125101
A/Accession: T18528
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1253 <TOW>
A/Cross-references: EMBL:L27221; NID:G960281; PID:G960283; PIDN:AAA74894.1
A/Description: may be involved in electron transfer from pyruvate to flavodoxin
C/Superfamily: pyruvate (flavodoxin) dehydrogenase; ferredoxin [4Fe-4S] homology
C/Keywords: 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein; oxidoredu

Query Match

Best local similarity 77.8%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVAXSLF 9

DB 906 GIACSLF 914

RESULT 2

A87066
conserved hypothetical protein ML1255 (imported) - Mycobacterium leprae
C/Species: Mycobacterium leprae
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C/Accession: A87066
R/Cole, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.;
R.; Davies, R.M.; Devlin, K.; Duthey, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holtro
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.;
A/Title: Massive gene decay in the leprosy bacillus.
A/Reference number: A86909; MUID:21128732; PMID:11234002
A/Accession: A87066
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-153 <STO>
A/Cross-references: GB:AL450380; NID:G13093193; PIDN:CMC31636.1; GSPDB:GN00147
C/Genetics:
A/Genes: ML1255
C/Superfamily: Mycobacterium leprae hypothetical protein MLCB1610.16

Query Match 83.8%; Score 31; DB 2; Length 163;
 Best Local Similarity 77.8%; Pred. No. 7.4;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 DB 132 GVAASSAF 140

RESULT 3

C70866
 hypothetical protein RV2468c - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 28-Jul-2000
 C/Accession: C70866
 R/Cole, S.T.; Broese, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A/Authors: Sgares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500; MUID:98295987; PMID:9634230
 A/Accession: C70866
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-167 <COL>
 A/Cross-references: GB:AL021246; GB:AL123456; NID:G3261507; PIDN:CAA16045.1; PID:G279150
 A/Experimental source: strain H37RV
 C/Genetics:
 A/Gene: RV2468c
 C/Superfamily: Mycobacterium leprae hypothetical protein MLCB1610.16

Query Match 83.8%; Score 31; DB 2; Length 167;
 Best Local Similarity 77.8%; Pred. No. 7.6;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 DB 136 GVAASSAF 144

RESULT 4

T15243
 hypothetical protein MLCB1610.16 [imported] - Mycobacterium leprae

C/Species: Mycobacterium leprae
 C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 28-Jul-2000
 C/Accession: T15243
 R/James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 A/Submitted to the EMBL Data Library, March 1999
 A/Reference number: Z22949
 A/Accession: T15243
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-169 <JAM>
 A/Cross-references: EMBL:AL049913; PIDN:CA843162.1
 A/Experimental source: cosmid B1610
 C/Genetics:
 A/Note: MLCB1610.16
 C/Superfamily: Mycobacterium leprae hypothetical protein MLCB1610.16

Query Match 83.8%; Score 31; DB 2; Length 169;
 Best Local Similarity 77.8%; Pred. No. 7.7;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 DB 138 GVAASSAF 146

RESULT 5

T13566
 translocation lipoprotein J - Yersinia pestis plasmid pCD1
 C/Species: Yersinia pestis

C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
 C/Accession: T43566; T42861
 R/Hu, P.; Elliott, J.; McCreedy, P.; Skowronski, E.; Garnes, J.; Kobayashi, A.; Bruba, J.; Bacteriol. 180, 5192-5202, 1998
 A/Title: Structural organization of virulence-associated plasmids of Yersinia pestis.
 A/Reference number: Z22578; MUID:98422474; PMID:9748454
 A/Accession: T43566
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-244 <HUP>
 A/Cross-references: EMBL:AF03946; NID:G2996222; PIDN:AAC62607.1; PID:G2996284
 A/Experimental source: strain KIM
 R/Barry, R.D.; Straley, S.C.; Fehreston, J.D.; Rose, D.J.; Gregor, J.; Blattner, F.R.
 Infect. Immun. 66, 4611-4623, 1998
 A/Title: DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of Yersinia
 A/Reference number: Z22273; MUID:98427122; PMID:9746557
 A/Accession: T42861
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-244 <EBR>
 A/Cross-references: EMBL:AF074612; NID:G3822037; PIDN:AAC69775.1; PID:G3822055
 A/Experimental source: strain KIMS
 C/Genetics:
 A/Gene: yscJ
 A/Genome: yscJ
 C/Superfamily: invasion secretion factor mxJ

Query Match 81.1%; Score 30; DB 2; Length 244;
 Best Local Similarity 66.7%; Pred. No. 19;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 DB 143 GVAASAVF 151

RESULT 6

A40049
 virC-region lipoprotein yscJ (YIPB) - Yersinia enterocolitica plasmid pYV

C/Species: Yersinia enterocolitica
 C/Date: 21-Feb-1992 #sequence_revision 21-Feb-1992 #text_change 17-Mar-2000
 C/Accession: A40049
 R/Michels, T.; Vancoteghem, J.C.; Lambert de Rouvroit, C.; Chini, B.; Gustin, A.; Bo, J.; Bacteriol. 173, 4994-5009, 1991
 A/Title: Analysis of virC, an operon involved in the secretion of Yop proteins by Yersinia
 A/Reference number: A40361; MUID:91317716; PMID:1860816
 A/Accession: A40049
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-244 <MIC>
 A/Cross-references: GB:M74011; NID:G155549; PIDN:AAC37027.1; PID:G155559
 C/Genetics:
 A/Genome: plasmid
 C/Superfamily: invasion secretion factor mxJ
 C/Keywords: lipid binding; lipoprotein

Query Match 81.1%; Score 30; DB 2; Length 244;
 Best Local Similarity 66.7%; Pred. No. 19;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 DB 143 GVAASAVF 151

RESULT 7

B83602
 conserved hypothetical protein PA0340 [imported] - Pseudomonas aeruginosa (strain PAC)

C/Species: Pseudomonas aeruginosa
 C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 01-Mar-2002
 C/Accession: B83602
 R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L

.: Loty, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A/Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
 A/Reference number: AB2950; MUID:20437337; PMID:10984043
 A/Accession: B83602
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-267 <STO>
 A/Cross-references: GB:AE004472; GB:AE004091; NID:g9946188; PIDN:AA03729.1; GSPDB:GN001
 A/Experimental source: strain PA01
 A/Genetics:
 A/Gene: PA0340
 C/Superfamily: hypothetical protein H1092

Query Match 81.1%; Score 30; DB 2; Length 267;
 Best Local Similarity 66.5%; Pred. No. 21;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 Db 219 GVAATSLF 227

RESULT 8
 AC2730
 conserved hypothetical protein Atu1248 [imported] - *Agrobacterium tumefaciens* (strain C5
 C/Species: *Agrobacterium tumefaciens*
 C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C/Accession: AC2730
 R/Wood, D.W.; Secubal, U.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
 exage, P.; Gillet, W.; Grant, C.; Genthner, D.; Kutyavyn, T.; Levy, R.; Li, M.; McCell
 /Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
 A/Reference number: AB2577; MUID:21608550; PMID:11743193
 A/Accession: AC2730
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-305 <KUR>
 A/Cross-references: GB:AE006688; PIDN:AA142257.1; PID:g17739654; GSPDB:GN00186
 A/Experimental source: strain C58 (Dugont)
 C/Genetics:
 A/Gene: Atu1248
 A/Map position: circular chromosome

Query Match 81.1%; Score 30; DB 2; Length 305;
 Best Local Similarity 87.5%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VAAAXSLF 9
 Db 127 VAAATSLF 134

RESULT 9
 D97511
 probable integral membrane protein. (Al132674) [imported] - *Agrobacterium tumefaciens* (S
 C/Species: *Agrobacterium tumefaciens*
 C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
 C/Accession: D97511
 R/Gooder, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A/Title: Genome sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*
 A/Reference number: A97359; MUID:21608551; PMID:11743194
 A/Accession: D97511
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-305 <KUR>
 A/Cross-references: GB:AE007869; PIDN:AAK7045.1; PID:g15156295; GSPDB:GN00169
 C/Genetics:

A/Gene: AGR_C_2304
 A/Map position: circular chromosome

Query Match 81.1%; Score 30; DB 2; Length 305;
 Best Local Similarity 87.5%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VAAAXSLF 9
 Db 127 VAAATSLF 134

RESULT 10
 A29923
 carboxylesterase (EC 3.1.1.1), TCCD-induced 60K microsomal - rabbit
 C/Species: *Oryctolagus cuniculus* (domestic rabbit)
 C/Date: 08-Dec-1988 #sequence_revision 08-Dec-1988 #text_change 21-Aug-1998
 C/Accession: A29923; A29471
 R/Korza, G.; Ozols, J.
 J. Biol. Chem. 263, 3486-3495, 1988
 A/Title: Complete covalent structure of 60-kDa esterase isolated from 2,3,7,8-tetrach
 A/Reference number: A29923; MUID:86139431; PMID:3343253
 A/Accession: A29923
 A/Molecule type: protein
 A/Residues: 1-539 <KOR>
 R/Ozols, J.
 J. Biol. Chem. 262, 15316-15321, 1987

A/Title: Isolation and characterization of a 60-kilodalton glycoprotein esterase from
 A/Reference number: A29471; MUID:86033124; PMID:3667634
 A/Accession: A29471
 A/Molecule type: protein
 A/Residues: 1-71;193-208;436-446;532-539 <OZO>
 C/Superfamily: cholinesterase; cholinesterase homology
 C/Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein
 F/32-526/Domain: cholinesterase homology <CHE>
 F/61;363/Binding site: carbonyl (asn) (covalent) #status experimental
 F/195;441/Active site: Ser, His #status experimental

Query Match 81.1%; Score 30; DB 2; Length 539;
 Best Local Similarity 77.8%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 Db 222 GVAATSLF 230

RESULT 11
 B64399
 hypothetical protein MJ0797 - *Methanococcus jannaschii*
 C/Species: *Methanococcus jannaschii*
 C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C/Accession: B64399
 R/Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
 /son, U.D.; Sadow, P.W.; Hama, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
 A/Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannasc*
 A/Reference number: A64300; MUID:96337999; PMID:8688087
 A/Accession: B64399
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-367 <BLU>
 A/Cross-references: GB:U67524; GB:L77117; NID:g2826321; PIDN:AA898792.1; PID:g1591494
 A/Genetics:
 A/Map position: FOR720661-721764

Query Match 78.4%; Score 29; DB 2; Length 367;
 Best Local Similarity 87.5%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVAAXSLF 8

Db 27 GVAATSL 34

RESULT 12

hypothetical protein AT4G04480 [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
 C/Accession: F85056
 R/Anonymous: The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
 Nature 402, 769-777, 1999
 A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A/Reference number: A85001; MUID:20083488; PMID:10617198
 A/Accession: F85056
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-398 <STO>
 A/Cross-references: GB:NC_001268; NID:g7267205; PIDN:CA877916.1; GSPDB:GN00140
 C/Genetics:
 A/Gene: AT4G04480
 A/Map position: 4

Query Match 78.4%; Score 29; DB 2; Length 398;
 Best Local Similarity 66.7%; Pred. No. 52;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 |||
 Db 289 GVSAYSVF 297

RESULT 13

cobalamin biosynthesis protein N - Methanobacterium thermoautotrophicum (strain Delta H)
 C/Species: Methanobacterium thermoautotrophicum
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C/Accession: C69224
 R/Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 ; Qiu, D.; Spedifora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
 X.; S.; Chutich, G.M.; Daniels, C.U.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funcit
 A/Reference number: A69000; MUID:98037514; PMID:9371463
 A/Accession: C69224
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-1668 <MTH>
 A/Cross-references: GB:AE000868; GB:AE000666; NID:g2622025; PIDN:AAB85426.1; PID:g262202
 A/Experimental source: strain Delta H
 C/Genetics:
 A/Gene: MTH928
 A/Start codon: GTG
 C/Superfamily: Methanobacterium thermoautotrophicum cobalamin biosynthesis protein N

Query Match 78.4%; Score 29; DB 1; Length 1668;
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 |||
 Db 1273 GVAATSLF 1281

RESULT 14

aldolase C - chicken (fragment)
 C/Species: Gallus gallus (chicken)
 C/Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 22-Jun-1999
 C/Accession: I51291
 R/Welshman-Mantha, R.L.; Toljan, D.R.
 J. Cell. Biochem. 57, 423-431, 1995
 A/Title: Noncoordinate changes in the steady-state mRNA expressed from aldolase A and a

A/Reference number: I51291; MUID:95286677; PMID:7768978
 A/Accession: I51291
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-42 <MET>
 A/Cross-references: GB:S88288; NID:g999389; PIDN:AAB34479.1; PID:g999390
 C/Superfamily: Fructose-bisphosphate aldolase

Query Match 75.7%; Score 28; DB 2; Length 42;
 Best Local Similarity 66.7%; Pred. No. 10;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 |||
 Db 28 GVAATSLF 36

RESULT 15

Agd protein [imported] - Staphylococcus aureus (strain N315)
 C/Species: Staphylococcus aureus
 C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C/Accession: C89995
 R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; C
 ma, A.; Mitutani, U.; Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsuku, K.
 Lancet 357, 1225-1240, 2001
 A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A/Reference number: A89758; MUID:21311952; PMID:11418146
 A/Accession: C89995
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-47 <KTR>
 A/Cross-references: GB:BA000018; PID:g13701831; PIDN:BAE43124.1; GSPDB:GN00149
 A/Experimental source: strain N315
 C/Genetics:
 A/Gene: agdD

Query Match 75.7%; Score 28; DB 2; Length 47;
 Best Local Similarity 77.8%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 |||
 Db 24 GVAACSLF 32

Search completed: January 29, 2004, 14:58:21
 Job time : 12.1176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:45:44 ; Search time 5.8253 Seconds
(without alignments)
72.678 Million cell updates/sec

Title: US-10-032-950-3

Sequence: 1 GYAXSLP 9

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	30	81.1	244 1 YSCJ_YEREN	Q01251 yersinia en
2	30	81.1	244 1 YSCJ_YERPE	Q00926 yersinia pe
3	30	81.1	539 1 ESI1_RABIT	P12337 oryctolagus
4	29	78.4	367 1 Y797_METUA	Q56207 methanococc
5	29	78.4	566 1 HCP_FUSIN	Q84611 fusobacteri
6	28	75.7	282 1 SRC2_CAEL	P45571 caenobacteri
7	28	75.7	330 1 PDXA_XYLPA	Q95639 xyliella fas
8	28	75.7	331 1 ADT1_WHEAT	Q41629 triticum ae
9	28	75.7	331 1 ADT2_WHEAT	Q41630 triticum ae
10	28	75.7	338 1 SFUA_SERMA	P21408 serratia ma
11	28	75.7	363 1 ALFA_HUMAN	P04075 homo sapien
12	28	75.7	363 1 ALFA_MOUSE	P05064 mus musculu
13	28	75.7	363 1 ALFA_MOUSE	P00883 oryctolagus
14	28	75.7	363 1 ALFA_MOUSE	P05065 rattus norv
15	28	75.7	363 1 ALFA_MOUSE	Q91957 mus musculu
16	28	75.7	362 1 ADT1_ORYSA	P31691 oryza sativ
17	28	75.7	366 1 ADT1_SOLTA	P25083 solanum tub
18	28	75.7	387 1 ADT1_MAIZE	P04709 zea mays (m
19	28	75.7	387 1 ADT2_MAIZE	P12857 zea mays (m
20	28	75.7	442 1 MNTH_ANASP	Q82850 anabaena sp
21	28	75.7	522 1 SINI_CHICK	Q96653 gallus gall
22	27	73.0	283 1 PRIT_TRIAL	P20015 trititachiu
23	27	73.0	325 1 YU40_YEAST	P47120 saccharomyc
24	27	73.0	342 1 Y762_METUA	Q58172 methanococc
25	27	73.0	363 1 ALF1_LAMUA	P53445 lampetra ja
26	27	73.0	406 1 LMPI_MOUSE	P11438 mus musculu
27	27	73.0	407 1 LMPI_MOUSE	P14562 rattus norv
28	27	73.0	410 1 NOB1_NEIMA	Q91049 neisseria m
29	27	73.0	410 1 NOB1_NEIMA	Q91049 neisseria m
30	27	73.0	451 1 YCJ1_ECOLI	P76637 escherichia
31	27	73.0	673 1 Y552_HUMAN	Q60293 homo sapien
32	27	73.0	953 1 B3A4_RAT	Q84572 rattus norv
33	27	73.0	955 1 B3A4_RABIT	Q93471 oryctolagus

ALIGNMENTS

RESULT 1	ID	YSCJ_YEREN	STANDARD	PRT	244 AA
34	27	73.0	983 1	B3A4_HUMAN	Q96931 homo sapien
35	27	73.0	1321 1	AB11_RABIT	Q91043 oryctolagus
36	26	70.3	79 1	Y197_CAEL	Q65511 caenobacteri
37	26	70.3	179 1	YCFU_ECOLI	P37796 escherichia
38	26	70.3	213 1	THIE_THENC	Q91474 thermoplas
39	26	70.3	220 1	THIE_AGR5	Q84684 agrobacteri
40	26	70.3	246 1	TRUA_THERO	Q97974 thermoplas
41	26	70.3	247 1	CYCX_DESDE	P81040 desulfovibr
42	26	70.3	280 1	KIC3_XENLA	P05782 xenopus lae
43	26	70.3	318 1	KDGT_ERWCA	Q95522 erwinia car
44	26	70.3	362 1	Y022_METUA	Q60342 methanococc
45	26	70.3	364 1	ALF2_LAMUA	P53446 lampetra ja

Query Match: 61.1%; Score 30; DB 1; Length 244;
Best Local Similarity: 66.7%; Pred. No. 8.9; 1; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 1;

QY 1 GVAAXSLF 9
 DB 143 GVAASASVF 151

RESULT 2
 YSCU_YERPE STANDARD; PRT; 244 AA.
 AC Q00926;
 ID 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE YOP proteins translocation lipoprotein J precursor (Lipoprotein Y1p3)
 DE (Low calcium response locus protein Ka).
 GN YSCU OR YLPB OR LCRA OR YPCD1.59 OR Y5019 OR Y0022.
 OS Yersinia pestis, and
 OS Yersinia pseudotuberculosis.
 OC Plasmid pCD1, and plasmid pIB1.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OC NCBI_TaxId=632, 633;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Y.pestis; STRAIN=KIMS / Biovar Mediaevalis; PLASMID=pCD1;
 RX MEDLINE=98427122; PubMed=9746557;
 RA Perry R.D., Straley S.C., Fetherston J.D., Rose D.J., Gregor J.,
 RA Blattner F.R.;
 RT "DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of
 RT Yersinia pestis KIMS.";
 RL Infect. Immun. 66:4611-4623(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Y.pestis; STRAIN=KIMS / Biovar Mediaevalis; PLASMID=pCD1;
 RX MEDLINE=9842474; PubMed=9748454;
 RA Hu P., Elliott J., McCreedy P., Skowronski E., Garnea J.,
 RA Kobayashi A., Brubaker R.R., Garcia E.;
 RT "Structural organization of virulence-associated plasmids of Yersinia
 RT pestis.";
 RL J. Bacteriol. 180:5192-5202(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Y.pestis; STRAIN=CO-92 / Biovar Orientalis; PLASMID=pCD1;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Tibball R.W., Holden M.T.G.,
 RA Prentice M.B., Sepahinia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Felwell T., Hamlin N., Holtroyd S., Jagers K., Kariyasek A.V.,
 RA Leather S., Kule S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skellton J., Stevens K., Whitehead S., Barrell B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Y.pseudotuberculosis; STRAIN=YPIII; PLASMID=pIB1;
 RX MEDLINE=92250432; PubMed=1577700;
 RA Rimplaenen M., Forsberg A., Wolf-Matz H.;
 RT "A novel protein, Icro, involved in the low-calcium response of
 RT Yersinia pseudotuberculosis shows extensive homology to YopH.";
 RL J. Bacteriol. 174:3355-3363(1992).
 CC -1- FUNCTION: REQUIRED FOR THE EXPORT PROCESS OF THE YOP PROTEINS.
 CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
 CC anchor.
 CC -1- INDUCTION: At 37 degrees Celsius in the absence of calcium.
 CC -1- SIMILARITY: BELONGS TO THE YSCU FAMILY OF LIPOPROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: AF074612; AAC69775.1;
 CC EMBL: AF053946; AAC62607.1;
 CC EMBL: AL117189; CAB54936.1;
 CC EMBL: M83986; AAA27651.1;
 CC PIR: T43566; T43566.
 CC InterPro: IPR002282; SecITIMPK.
 CC InterPro: IPR006182; YscU_F11F.
 CC Pfam: PF01514; YscU_F11F.1.
 CC PRINTS: PR01339; TYF30MKRPOT.
 CC PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 CC Transprot: Protein transport; Outer membrane; Transmembrane;
 CC Lipoprotein; Signal; Plasmid; Virulence; Complete proteome.
 CC SIGNAL 1 18 PROBABLE.
 CC CHAIN 19 244 YOP PROTEINS TRANSLOCATION LIPOPROTEIN J.
 CC LIPID 19 19 N-ACYL DIGLYCERIDE (POTENTIAL).
 CC TRANSMEM 218 238 POTENTIAL.
 CC SEQUENCE 244 AA; 27042 MW; 8DDA9E712938D8BC CRC64;
 SQ

Query Match 81.1%; Score 30; DB 1; Length 244;
 Best Local Similarity 66.7%; Pred. No. 8.9;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0.

QY 1 GVAAXSLF 9
 DB 143 GVAASASVF 151

RESULT 3
 EST1_RABIT STANDARD; PRT; 539 AA.
 AC P12337;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Liver carboxylesterase 1 (EC 3.1.1.1).
 OS Oryctolagus cuniculus (Rabbit).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 CC NCBI_TaxId=9986;
 OX [1]
 RP SEQUENCE.
 RX MEDLINE=86139431; PubMed=3343253;
 RA Korza G., Ozols J.;
 RT "Complete covalent structure of 60-kDa esterase isolated from
 RT 2,3,7,8-tetrachlorodibenzo-p-dioxin-induced rabbit liver
 RT microsomes.";
 RL J. Biol. Chem. 263:3486-3495(1988).
 RN [2]
 RP SEQUENCE OF 1-70 AND 532-539.
 RX MEDLINE=88031124; PubMed=3667634;
 RA Ozols J.;
 RT "Isolation and characterization of a 60-kilodalton glycoprotein
 RT esterase from liver microsomal membranes.";
 RL J. Biol. Chem. 262:15316-15321(1987).
 CC -1- FUNCTION: INVOLVED IN THE DETOXIFICATION OF XENOBIOTICS AND IN
 CC THE ACTIVATION OF ESTER AND AMIDE PRODUCTS.
 CC -1- CATALYTIC ACTIVITY: A carboxylic ester + H(2)O = an alcohol + a
 CC carboxylic anion.
 CC -1- SUBUNIT: Monomer.
 CC -1- SUBCELLULAR LOCATION: Microsomal membrane, lumen of endoplasmic
 CC reticulum.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
 CC PIR: A29923; A29923.
 CC HSP: P21836; IMAH.
 CC InterPro: IPR002018; CarbesteraeB.
 CC InterPro: IPR000379; Ser_estr_ site.
 CC Pfam: PF00135; Coesterase_1.
 CC PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
 CC PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
 CC Glycoprotein; Hydrolase; Serine esterase; Endoplasmic reticulum.
 CC ACT_SITE 195 195 BY SIMILARITY.
 CC ACT_SITE 441 441 BY SIMILARITY.


```

QY      1 GVAAXSLF 9
DB      43 GVAAXSLF 51

RESULT 6
SRG2_CAEEL STANDARD; PRT; 282 AA.
ID      SRG2_CAEEL
AC      P46571;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DE      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Serenine receptor class gamma 2 (Srg-2 protein).
GN      Srg-2 OR C18F10.5.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC      Rhabditidae; Peloderinae; Caenorhabditis.
OX      NCBI_TaxID=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RA      Latteille P.;
RL      Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC      -1- SIMILARITY: BELONGS TO THE C.ELEGANS RECEPTOR-LIKE PROTEIN SRG
CC      FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U00049; AAC47052.1; -
DR      PIR; T15556; T15556.
DR      WormPep; C18F10.5; CE00782.
DR      InterPro; IPR00609; Srg.
DR      Pfam; PF02118; Srg; 1.
DR      PRINTS; PR00698; TMPTROTEINSRG.
KM      Transmembrane; Multigene family.
FT      TRANSMEM 25 45
FT      TRANSMEM 112 132 POTENTIAL.
FT      TRANSMEM 151 171 POTENTIAL.
FT      TRANSMEM 199 219 POTENTIAL.
FT      TRANSMEM 232 252 POTENTIAL.
SQ      SEQUENCE 282 AA; 32675 MW; 4FD6AD60C65FD27 CRC64;

Query Match      75.7%; Score 28; DB 1; Length 282;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

RP      SEQUENCE FROM N.A.
RC      STRAIN=9a5c;
RX      MEDLINE=20365717; PubMed=10910347;
RA      Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA      Barroes M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
RA      Bueno M.R.P., Camargo A.A., Camargo J.E.A., Carraro D.M., Carter H.,
RA      Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA      Coutinho L.B., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA      Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA      Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA      Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA      Ho P., Honeisel U.D., Junqueira M.L., Kemper E.L., Kitajima U.P.,
RA      Krieger J.E., Kuramae E.B., Laigret F., Lamais M.R., Leite L.C.C.,
RA      Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA      Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
RA      Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA      Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA      Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA      Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA      de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
RA      Pelxoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,
RA      Queiroz R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA      de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawaaki H.B.,
RA      da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A., Jr.,
RA      da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA      de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,
RA      Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA      Zago M.A., Zatz M., Zeldanis J., Zetudal J.C.;
RL      "The genome sequence of the plant pathogen Xylella fastidiosa,"
RL      Nature 406:151-159(2000).
CC      -1- FUNCTION: Catalyzes the NAD-dependent oxidation of 4-
CC      (phosphohydroxy)-L-threonine (HTP) into 3-amino-3-oxo-4-
CC      (phosphohydroxy)butyric acid which spontaneously decarboxylate to
CC      form 1-amino-3- (phosphohydroxy)propan-2-one (3-amino-2-oxopropyl
CC      phosphate) (By similarity).
CC      -1- CATALYTIC ACTIVITY: 4 (phosphonooxy)-threonine + NAD(+) = 2-
CC      amino-3-oxo-4-phosphonooxybutyrate + NADH.
CC      -1- PATHWAY: De novo synthesis of pyridoxine (vitamin B6) and
CC      pyridoxal phosphate.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC      -1- SIMILARITY: BELONGS TO THE PDXA FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AB003923; AAF83649.1; -
DR      PIR; D82756; D82756.
DR      HAMAP; MF_00536; -; 1.
DR      InterPro; IPR005255; PdxA.
DR      Pfam; PF04156; PdxA; 1.
DR      TIGRFAMs; TIGR00837; PdxA; 1.
KM      Pyridoxine biosynthesis; Oxidoreductase, NAD, Complete proteome.
SQ      SEQUENCE 330 AA; 34553 MW; A981237FC844CF71 CRC64;

Query Match      75.7%; Score 28; DB 1; Length 330;
Best Local Similarity 77.8%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ADP/ATP carrier protein 1, mitochondrial precursor (ADP/ATP
DE translocase 1) (Adenine nucleotide translocator 1) (ANT 1).
GN ANT-G1.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxId=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Iacobazzi V., Poli A., Bianco A., Palmieri F.;
RT "Nucleotide sequences of two genes (ANT-G1 and ANT-G2) encoding the
RT adenine nucleotide translocator of wheat mitochondria.";
RL (In) Plant Gene Register PGR96-016.
CC -1- FUNCTION: CATALYZES THE EXCHANGE OF ADP AND ATP ACROSS THE
CC MITOCHONDRIAL INNER MEMBRANE.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: Belongs to the mitochondrial carrier family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X95863; CA65119.1; -
CC InterPro; IPR002067; Mlt_carrier.
CC InterPro; IPR001993; Mitoch_carrier.
CC Pfam; PF00153; mltc_carr; 3.
CC PRINTS; PR00926; MITOCARRIER.
CC PROSITE; PS00215; MITOCH_CARRIER; 2.
CC Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
CC Transist peptide; Multigene family.
CC TRANSIT 1 ?
CC CHAIN ? 331 ADP/ATP CARRIER PROTEIN 1.
CC TRANSMEM 35 52 1 (POTENTIAL).
CC TRANSMEM 97 115 2 (POTENTIAL).
CC TRANSMEM 140 157 3 (POTENTIAL).
CC TRANSMEM 201 220 4 (POTENTIAL).
CC TRANSMEM 240 257 5 (POTENTIAL).
CC TRANSMEM 296 314 6 (POTENTIAL).
CC SEQUENCE 331 AA; 35961 MW; D0C9464DD43AEAF3 CRC64;
SQ

```

```

Query Match 75.7%; Score 28; DB 1; Length 331;
Best Local Similarity 66.7%; Pred. NO. 36;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 GVAAXSLF 9
DB 144 GAAGASSLF 152

```

```

RESULT 9
ADT2_WHEAT STANDARD; PRT; 331 AA.
AC ADT2_WHEAT
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ADP/ATP carrier protein 2, mitochondrial precursor (ADP/ATP
DE translocase 2) (Adenine nucleotide translocator 2) (ANT 2).
GN ANT-G2.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxId=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Iacobazzi V., Poli A., Bianco A., Palmieri F.;
RT "Nucleotide sequences of two genes (ANT-G1 and ANT-G2) encoding the
RT adenine nucleotide translocator of wheat mitochondria.";
RL (In) Plant Gene Register PGR96-016.
CC -1- FUNCTION: CATALYZES THE EXCHANGE OF ADP AND ATP ACROSS THE
CC MITOCHONDRIAL INNER MEMBRANE.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: Belongs to the mitochondrial carrier family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X95864; CA65120.1; -
CC InterPro; IPR002067; Mlt_carrier.
CC InterPro; IPR001993; Mitoch_carrier.
CC Pfam; PF00153; mltc_carr; 3.
CC PRINTS; PR00926; MITOCARRIER.
CC PROSITE; PS00215; MITOCH_CARRIER; 2.
CC Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport;
CC Transist peptide; Multigene family.
CC TRANSIT 1 ?
CC CHAIN ? 331 ADP/ATP CARRIER PROTEIN 2.
CC TRANSMEM 35 52 1 (POTENTIAL).
CC TRANSMEM 97 115 2 (POTENTIAL).
CC TRANSMEM 140 157 3 (POTENTIAL).
CC TRANSMEM 201 220 4 (POTENTIAL).
CC TRANSMEM 240 257 5 (POTENTIAL).
CC TRANSMEM 296 314 6 (POTENTIAL).
CC SEQUENCE 331 AA; 35790 MW; 265DE6AA41557908 CRC64;
SQ

```

```

Query Match 75.7%; Score 28; DB 1; Length 331;
Best Local Similarity 66.7%; Pred. NO. 36;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 GVAAXSLF 9
DB 144 GAAGASSLF 152

```

```

RESULT 10
SFUA_SERMA STANDARD; PRT; 338 AA.
ID SFUA_SERMA
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Iron(III)-binding periplasmic protein precursor.
GN SFUA.
OS Serratia marcescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Serratia.
OX NCBI_TaxId=615;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=90130286; PubMed=2404942;
RA Angerer A., Gaisner S., Braun V.;
RT "Nucleotide sequences of the sfuA, sfuB, and sfuC genes of Serratia
RT marcescens suggest a periplasmic-binding-protein-dependent iron
RT transport mechanism.";

```

```

RL J. Bacteriol. 172:572-578(1990).
CC -1- FUNCTION: INVOLVED IN A PERIPLASMIC BINDING-PROTEIN-DEPENDENT
CC IRON TRANSPORT SYSTEM, WHICH IS INDEPENDENT OF TONB, EXB,
CC SIDEROPHORE, OR ANY OTHER MEMBRANE PROTEIN. HOWEVER, IT IS NOT
CC CLEAR HOW FE(III) IS SOLUBILIZED AND TAKEN UP ACROSS THE OUTER
CC MEMBRANE.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
CC PROTEIN FAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, M33815; AAA26573.1; -.
DR PIR, A35108; QRSUA.
DR HSSP, P35755; IMRP.
DR InterPro; IPR006059; SBP_dac.1.
DR InterPro; IPR006061; SBP_dom1.
DR Pfam; PF01547; SBP_dac.1; 1.
DR PROSITE; PS01037; SBP_BACTERIAL.1; 1.
KW Transport; Iron transport; Periplasmic; Signal.
FT SIGNAL
FT CHAIN
FT SIGNAL
SQ SEQUENCE 338 AA; 36157 MW; DDF5FAA52301A716 CRC64;

Query Match
Best Local Similarity 75.7%; Score 28; DB 1; Length 338;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
Db 126 GIARSTVF 134

RESULT 11
ALFA_HUMAN STANDAR; PRT; 363 AA.
ID ALFA_HUMAN
AC P04075;
DT 01-NOV-1986 (Rel. 03; Created)
DT 01-APR-1988 (Rel. 07; Last sequence update)
DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE Fructose-bisphosphate aldolase A (EC 4.1.2.13) (muscle-type aldolase)
DE (lung cancer antigen NY-LU-1).
GN ALDOA OR ALDA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
XP [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RA MEDLINE=87161904; PubMed=3030757;
RA Izzo P., Costanzo P., Lupco A., Ridpa E., Borghese A.M., Paolella G.,
RA Salvatore F.;
RA "A new human species of aldolase A mRNA from fibroblasts.";
RT Eur. J. Biochem. 164:9-13(1987).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=88271327; PubMed=3391172;
RA Izzo P., Costanzo P., Lupco A., Ridpa E., Paolella G., Salvatore F.;
RT "Human aldolase A gene. Structural organization and tissue-specific
RT expression by multiple promoters and alternate mRNA processing.";
RL Eur. J. Biochem. 174:565-578(1988).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=85306986; PubMed=3840020;
RA Sakibara M., Mukai T., Hori K.;

RT "Nucleotide sequence of a cDNA clone for human aldolase: a messenger
RT RNA in the liver.";
RL Biochem. Biophys. Res. Commun. 134:413-420(1985).
RN [4]
RP SEQUENCE FROM N.A.
RA MEDLINE=9153319; PubMed=1999195;
RA Mukai T., Arai Y., Yatsuki H., Joh K., Hori K.;
RT "An additional promoter functions in the human aldolase A gene, but
RT not in rat.";
RL Eur. J. Biochem. 195:781-787(1991).
RN [5]
RP SEQUENCE
RA MEDLINE=88183272; PubMed=3355497;
RA Freeman P.S., Dunbar B., Fothergill Gilmore L.A.;
RT "The complete amino acid sequence of human skeletal-muscle fructose-
RT biphosphate aldolase.";
RL Biochem. J. 249:779-788(1988).
RN [6]
RP SEQUENCE OF 1-62 AND 147-357.
RA MEDLINE=84126818; PubMed=6696436;
RA Freeman P.S., Dunbar B., Fothergill L.A.;
RT "Human skeletal-muscle aldolase: N-terminal sequence analysis of
RT CNBR- and o-iodosobenzoic acid-cleavage fragments.";
RL Arch. Biochem. Biophys. 228:342-352(1984).
RN [7]
RP SEQUENCE OF 1-107 FROM N.A.
RA MEDLINE=88155643; PubMed=3441006;
RA Waite P., Gautron S., Hakim V., Gregori C., Mennecier F., Kahn A.;
RT "Characterization of three optional promoters in the 5' region of the
RT human aldolase A gene.";
RL J. Mol. Biol. 197:425-438(1987).
RN [8]
RP SEQUENCE OF 138-363 FROM N.A.
RA MEDLINE=88046782; PubMed=3674018;
RA Tolan D.R., Niclas J., Bruce B.D., Lebo R.V.;
RT "Evolutionary implications of the human aldolase-A, -B, -C, and -
RT pseudogene chromosome locations.";
RL Am. J. Hum. Genet. 41:907-924(1987).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RA MEDLINE=90242948; PubMed=2335208;
RA Gambin S.J., Cooper B., Millar J.R., Davies G.J., Littlechild J.A.,
RA Watson H.C.;
RT "The crystal structure of human muscle aldolase at 3.0-A resolution.";
RL FEBS Lett. 262:282-286(1990).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RA MEDLINE=91278081; PubMed=2056525;
RA Gambin S.J., Davies G.J., Grimes J.M., Jackson R.M.,
RA Littlechild J.A., Watson H.C.;
RT "Activity and specificity of human aldolases.";
RL J. Mol. Biol. 219:573-576(1991).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RA MEDLINE=99156067; PubMed=10048322;
RA Dalby A., Dauber Z., Littlechild J.A.;
RT "Crystal structure of human muscle aldolase complexed with fructose
RT 1,6-bisphosphate: mechanistic implications.";
RL Protein Sci. 8:291-297(1999).
RN [12]
RP VARIANT HEMOLYTIC ANEMIA GLY-128.
RA MEDLINE=88068641; PubMed=2825199;
RA Kishi H., Mukai T., Hirose A., Fujii H., Miwa S., Hori K.;
RT "Human aldolase A deficiency associated with a hemolytic anemia:
RT chemical aldolase due to a single base mutation.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8623-8627(1987).
RN [13]
RP VARIANT HEMOLYTIC ANEMIA GLY-128.
RA MEDLINE=91033340; PubMed=2228018;
RA Takasaki Y., Takahashi I., Mukai T., Hori K.;
RT "Human aldolase A of a hemolytic anemia patient with Asp-128-->Gly
RT substitution: characteristics of an enzyme generated in E. coli
RT transfected with the expression plasmid pHAD128G.";

```

```

FT STRAND 103 107
FT STRAND 112 114
FT TURN 116 117
FT TURN 122 124
FT TURN 128 129
FT HELIX 130 139
FT TURN 140 141
FT STRAND 144 151
FT HELIX 160 179
TURN 180 181
FT STRAND 183 190
FT HELIX 198 218
FT TURN 219 220
FT HELIX 223 225
FT STRAND 227 228
FT TURN 237 238
FT HELIX 245 257
FT TURN 258 259
FT TURN 262 263
FT STRAND 266 269
FT TURN 272 273
FT HELIX 276 288

Query Match 75.7%, Score 28, DB 1, Length 363;
Best Local Similarity 66.7%, Pred. No. 39;
Matches 6, Conservative 0, Mismatches 3, Indels 0, Gaps 0;

Oy 1 GVAAXSLF 9
Db 349 GAASESLF 357

RESULT 12
ID ALFA_MOUSE STANDARD; PRT; 363 AA.
AC POS054;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Fructose-bisphosphate aldolase A (EC 4.1.2.13) (Muscle-type aldolase)
DE (Aldolase 1).
GN ALDOA OR ALDO1.
OS Mus musculus (Mouse).
OC Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=880965598; PubMed=3697100;
RA Weste A., Stauffer J., Tolan D.R., Ciejek-Baez E.;
RT "Sequence of a mouse brain aldolase A cDNA.";
RL Nucleic Acids Res. 15:10595-10595 (1987).
RN [2]
RP SEQUENCE OF 1-265 AND 294-363 FROM N.A.
RX MEDLINE=90307699; PubMed=2365699;
RA Stauffer J.K., Colbert M.C., Ciejek-Baez E.;
RT "Nonconservative utilization of aldolase A alternative promoters.";
RL J. Biol. Chem. 265:11773-11782 (1990).
RN [3]
RP SEQUENCE OF 98-354 FROM N.A.
RX MEDLINE=86192445; PubMed=3009179;
RA Paoletti G., Buono P., Mancini P., Izzo P., Salvatore F.;
RT "Structure and expression of mouse aldolase genes. Brain-specific
aldolase C amino acid sequence is closely related to aldolase A.";
RL Eur. J. Biochem. 156:229-235 (1986).
-1- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycero-
phosphate + D-glyceraldehyde 3-phosphate.
-1- PATHWAY: Glycolysis; sixth step.
-1- SUBUNIT: Homotrimer.
-1- MISCELLANEOUS: In vertebrates, three forms of this ubiquitous
glycolytic enzyme are found, aldolase A in muscle, aldolase B in
liver and aldolase C in brain.

```

RT "Studies on the structure of rabbit muscle aldolase. Ordering of the
 RT tryptic peptides; sequence of 164 amino acid residues in the NH2-
 RT terminal B/CN peptide.";
 RL Arch. Biochem. Biophys. 166:347-357(1975).
 RN [4]
 RN SEQUENCE OF 173-200, AND REVISIONS.
 RX MEDLINE=80109133; PubMed=534504;
 RA Benfield P.A., Forcina B.G., Gibbons I., Perham R.N.;
 RA "Extended amino acid sequences around the active-site lysine residue
 RA of class-I fructose 1,6-bisphosphate aldolases from rabbit muscle,
 RA sturgeon muscle, trout muscle and ox liver";
 RL Biochem. J. 183:429-444(1979).
 RN [5]
 RN SEQUENCE OF 251-363, AND REVISION.
 RX MEDLINE=75145172; PubMed=1122142;
 RA Lai C.-Y.;
 RT "Studies on the structure of rabbit muscle aldolase. Determination of
 RT the primary structure of the COOH-terminal B/CN peptide; the complete
 RT sequence of the subunit polypeptide chain";
 RL Arch. Biochem. Biophys. 166:358-368(1975).
 RN [6]
 RN SEQUENCE OF 37-55 AND 349-363 FROM N.A.
 RX MEDLINE=83167564; PubMed=6687628;
 RA Putney S.D., Herlihy W.C., Schimmel P.R.;
 RA "A new tropomyosin T and CDNA clones for 13 different muscle proteins,
 RT found by shotgun sequencing.";
 RL Nature 302:718-721(1983).
 RN [7]
 RN ACTIVE SITE.
 RX MEDLINE=74163196; PubMed=4857186;
 RA Hartman F.C., Welch M.H.;
 RA "Identification of the histidyl residue of rabbit muscle aldolase
 RT alkylated by N-bromosuccinylsuccinylamine phosphate.";
 RL Biochem. Biophys. Res. Commun. 57:85-92(1974).
 RN [8]
 RN ACTIVE SITE.
 RX MEDLINE=76190154; PubMed=55453;
 RA Hartman F.C., Brown J.P.;
 RT "Affinity labeling of a previously undetected essential lysyl residue
 RT in class I fructose bisphosphate aldolase.";
 RL J. Biol. Chem. 251:3057-3062(1976).
 RN [9]
 RN SUBSTRATE-BINDING SITE.
 RX MEDLINE=80046697; PubMed=499203;
 RA Pethy L., Varadi A., Inesz U., Kovacs K.;
 RT "Identification of the C-1-phosphate-binding arginine residue of
 RT rabbit-muscle aldolase. Isolation of 1,2-cyclohexanedione-labeled
 RT peptide by chemisorption chromatography.";
 RL Eur. J. Biochem. 99:309-313(1979).
 RN [10]
 RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE=97143309; PubMed=8989120;
 RA Blom N., Szygusch U.;
 RT "Product binding and role of the C-terminal region in class I
 RT D-fructose 1,6-bisphosphate aldolase";
 RL Nat. Struct. Biol. 4:36-39(1997).
 CC -1- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glyceraldehyde
 CC phosphate + D-glyceraldehyde 3-phosphate.
 CC -1- PATHWAY: glycolysis; sixth step.
 CC -1- SUBUNIT: Tetramer of nearly identical chains, alpha and beta, which
 CC differ at only 1 position. N-360 in the alpha chain is denominated
 CC TO D-360 in the beta chain.
 CC -1- MISCELLANEOUS: In vertebrates, three forms of this ubiquitous
 CC glycolytic enzyme are found, aldolase A in muscle, aldolase B in
 CC liver and aldolase C in brain.
 CC -1- SIMILARITY: Belongs to class I fructose-bisphosphate aldolase
 CC family.
 CC -1- DATABASE: NAME=Worthington enzyme manual;
 CC WWW="http://www.worthington-biochem.com/manual/A/ALD.html".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: K02300; AAA31156.1; -
 DR EMBL: V00876; CAA24245.1; -
 DR EMBL: V00877; CAA24246.1; -
 DR PIR: A92444; ADABA.
 DR PDB: 1ADO; 24-DEC-97.
 DR PDB: 1EMD; 18-SEP-02.
 DR PDB: 1EWG; 18-SEP-02.
 DR PDB: 1EWG; 18-SEP-02.
 DR PDB: 1EX5; 18-SEP-02.
 DR PDB: 1JAE; 13-FEB-02.
 DR PDB: 6ALD; 05-JAN-00.
 DR InterPro: IPR00741; Aldolase_I.
 DR Pfam: PF00274; Glycolytic_enz_1.
 DR ProDom: PD001128; Aldolase_I; 1.
 DR PROSITE: PS00158; Aldolase_Class_I; 1.
 KM Lyase: Schiff base; Glycolysis; Multigene family; 3D-structure.
 FT INIT MET 0 0
 FT BINDING 55 55
 FT ACT_SITE 72 72
 FT ACT_SITE 107 107
 FT BINDING 146 146
 FT BINDING 229 229
 FT ACT_SITE 361 361
 FT ACT_SITE 363 363
 FT VARIANT 360 360
 FT CONFLICT 34 34
 FT CONFLICT 273 275
 FT CONFLICT 275 275
 FT CONFLICT 293 295
 FT CONFLICT 353 353
 FT HELIX 9 22
 FT TURN 23 23
 FT TURN 25 26
 FT STRAND 28 32
 FT HELIX 36 45
 FT TURN 46 47
 FT HELIX 52 63
 FT TURN 64 64
 FT HELIX 67 72
 FT STRAND 73 78
 FT HELIX 80 83
 FT TURN 84 84
 FT STRAND 86 86
 FT TURN 88 89
 FT STRAND 92 92
 FT HELIX 93 99
 FT TURN 100 101
 FT STRAND 103 107
 FT STRAND 112 114
 FT TURN 116 117
 FT STRAND 122 124
 FT TURN 128 129
 FT HELIX 130 139
 FT TURN 140 141
 FT STRAND 144 151
 FT HELIX 160 179
 FT TURN 180 181
 FT STRAND 183 190
 FT HELIX 198 218

C-1-PHOSPHATE GROUP OF THE SUBSTRATE.
 ESSENTIAL FOR CLEAVAGE OF THE SUBSTRATE
 TO GLYCERALDEHYDE 3-PHOSPHATE &
 DIHYDROXYACETONE PHOSPHATE.
 ESSENTIAL FOR CLEAVAGE OF THE SUBSTRATE
 TO GLYCERALDEHYDE 3-PHOSPHATE AND
 DIHYDROXYACETONE PHOSPHATE.
 C-1-PHOSPHATE GROUP OF THE SUBSTRATE,
 ALKYLATION INACTIVATES THE ENZYME.
 SCHIFF-BASE WITH DIHYDROXYACETONE-P.
 ALKYLATION INACTIVATES THE ENZYME.
 ESSENTIAL FOR THE SUBSEQUENT HYDROLYSIS
 OF THE DIHYDROXYACETONE SCHIFF BASE
 ESSENTIAL FOR ENHANCED ACTIVITY OF THE
 ENZYME TOWARD FRUCTOSE 1,6-BISPHOSPHATE
 AS COMPARED WITH FRUCTOSE 1-PHOSPHATE.
 N -> D (IN BETA CHAIN).
 E -> Q (IN REF. 3).
 G -> S (IN REF. 5).
 S -> E (IN REF. 5).
 KPM -> WPK (IN REF. 5).
 S -> R (IN REF. 6).

FT TURN 219 220
 FT HELIX 223 225
 FT STRAND 227 228
 FT TURN 237 238
 FT HELIX 245 259
 FT TURN 262 263
 FT STRAND 266 269
 FT TURN 272 273
 FT HELIX 276 288
 FT STRAND 296 301
 FT HELIX 303 313
 FT TURN 314 315
 FT HELIX 317 319
 FT HELIX 320 337
 FT TURN 328 340
 FT TURN 361 363

SQ SEQUENCE 363 AA; 39211 MW; 89CFCA31B181DE46 CRC64;
 Query Match 75.7%; Score 28; DB 1; Length 363;
 Best local similarity 66.7%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GYANXSLF 9
 Db 349 GAASESLF 357

RESULT 14
 ID ALFA_RAT STANDARD; PRT; 363 AA.
 AC P05065;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Fructose-bisphosphate aldolase A (EC 4.1.2.13) (muscle-type aldolase).
 GN ALDOA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;
 ON NCBI_Taxid=10116;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=86140113; PubMed=3753577;
 RX Mukai T., Joh K., Arai Y., Yatsuki H., Hori K.;
 RT "Tissue-specific expression of rat aldolase A mRNAs. Three molecular
 species differing only in the 5'-terminal sequences.";
 RL J. Biol. Chem. 261:3347-3354(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86083188; PubMed=2416636;
 RA Joh K., Mukai T., Yatsuki H., Hori K.;
 RT "Rat aldolase A messenger RNA: the nucleotide sequence and multiple
 mRNA species with different 5'-terminal regions.";
 RL Gene 39:17-24(1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87060996; PubMed=3783705;
 RA Joh K., Arai Y., Mukai T., Hori K.;
 RT "Expression of three mRNA species from a single rat aldolase A gene,
 differing in their 5' non-coding regions.";
 RL J. Mol. Biol. 190:401-410(1986).
 CC -I- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycerone
 phosphate + D-glyceraldehyde 3-phosphate.
 CC -I- PATHWAY: Glycolysis; sixth step.
 CC -I- SUBUNIT: Homotetramer.
 CC -I- MISCELLANEOUS: In vertebrates, three forms of this ubiquitous
 glycolytic enzyme are found, aldolase A in muscle, aldolase B in
 liver and aldolase C in brain.
 CC -I- SIMILARITY: Belongs to class I fructose-bisphosphate aldolase
 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M12919; AAA40714.1; -
 EMBL; M14420; AAA40715.1; -
 EMBL; X04261; CAA27815.1; -
 EMBL; X04262; CAA27815.1; JOINED.
 EMBL; X04263; CAA27815.1; JOINED.
 EMBL; X04264; CAA27815.1; JOINED.
 PIR; A24532; ADPRA.
 HSSP; P00863; IADO.
 InterPro; IPR00741; Aldolase_I.
 Pfam; PF00274; glycolytic_enz; 1.
 ProDom; PD001128; Aldolase_I; 1.
 PROSITE; PS00158; ALDOLASE_CLASS_I; 1.
 Lysase; Schiff base; Glycolysis; Multigene family.

INT MET 0
 BINDING 55 55 C-1-PHOSPHATE GROUP OF THE SUBSTRATE.
 BINDING 146 146 C-1-PHOSPHATE GROUP OF THE SUBSTRATE.
 BINDING 229 229 SCHIFF-BASE WITH DIHYDROXYACETONE-P.
 ACT_SITE 363 363 ESSENTIAL FOR ENHANCED ACTIVITY OF THE ENZYME TOWARD FRUCTOSE 1,6-BISPHOSPHATE AS COMPARED WITH FRUCTOSE 1-PHOSPHATE.

CONFLICT 144 144 F -> S (IN REF. 2).
 CONFLICT 164 164 M -> V (IN REF. 2).
 FT SEQUENCE 363 AA; 39220 MW; 48A0468B9E3B9DB8 CRC64;

Query Match 75.7%; Score 28; DB 1; Length 363;
 Best Local Similarity 66.7%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 DB 349 GAAATGSLF 357

RESULT 15
 ID ALPB_MOUSE STANDARD; PRT; 363 AA.
 AC Q91Y97; Q9K034; Q91W73;
 DT 15-SEP-2003 (Rel. 42; Created)
 DT 15-SEP-2003 (Rel. 42; Last sequence update)
 DT 15-SEP-2003 (Rel. 42; Last annotation update)
 DE Fructose-bisphosphate aldolase B (EC 4.1.2.13) (liver-type aldolase)
 DE (Aldolase 2).
 GN ALDOB OR ALDO2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SV.
 RA Funari V.A.; Tolan D.R.;
 RT "Mouse aldolase B (aldol2) genomic sequence of the open reading frame
 including first poly A site and signals."
 RT Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon, Kidney, and Liver;
 RA MEDLINE=2238257; PubMed=12477932;
 RA Strausberg R.D.; Feingold E.A.; Grouse L.H.; Derge J.G.;
 RA Klansner R.D.; Collins F.S.; Wagner L.; Shenmen C.M.; Schuler G.D.;
 RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
 RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.;
 RA Diachenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
 RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.;
 RA Brownstein W.U.; Usdin T.B.; Toshiyuki S.; Carninci P.; Prange C.;
 RA Raha S.S.; Loquellano N.A.; Peters G.J.; Adamson R.D.; Mullaly S.J.;
 RA Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;

Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.,
 Villalón D.K.; Muny D.M.; Sodergren E.V.; Lu X.; Gibbs R.A.,
 Whiting J.; Helton E.; Ketterman M.; Madan A.; Rodrigues S.; Sanchez A.,
 Blakeley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.,
 Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.,
 Butterfield V.S.N.; Krzywinski M.I.; Skalski U.; Smallos D.E.,
 Schermer A.; Schein J.E.; Jones S.J.W.; Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycerone
 phosphate + D-glyceraldehyde 3-phosphate.
 CC -1- PATHWAY: Glycolysis, sixth step.
 CC -1- SUBUNIT: Homotrimer (by similarity).
 CC -1- MISCELLANEOUS: In vertebrates, three forms of this ubiquitous
 glycolytic enzyme are found, aldolase A in muscle, aldolase B in
 liver and aldolase C in brain.
 CC -1- SIMILARITY: Belongs to class I fructose-bisphosphate aldolase
 family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; AF403567; AL06323.1; -
 DR EMBL; AF403565; AL06323.1; JOINED.
 DR EMBL; AF403566; AL06323.1; JOINED.
 DR EMBL; BC016435; AL016435.1; -
 DR EMBL; BC022113; AAH22113.1; -
 DR EMBL; BC024056; AAH24056.1; -
 DR EMBL; BC024112; AAH24112.1; -
 DR EMBL; BC026577; AAH26577.1; -
 DR EMBL; BC030724; AAH30724.1; -
 DR EMBL; BC030725; AAH30725.1; -
 DR EMBL; BC034169; AAH34169.1; -
 DR EMBL; BC034171; AAH34171.1; -
 DR EMBL; BC034172; AAH34172.1; -
 DR EMBL; BC034173; AAH34173.1; -
 DR EMBL; BC036130; AAH36130.1; -
 DR EMBL; BC036131; AAH36131.1; -
 DR EMBL; BC036132; AAH36132.1; -
 DR EMBL; BC036133; AAH36133.1; -
 DR InterPro; IPR000741; Aldolase_I.
 DR Pfam; PF00274; glycolytic_enz; 1.
 DR ProDom; PD001128; Aldolase_I; 1.
 DR PROSITE; PS00158; ALDOLASE_CLASS_I; 1.
 KW Lysase; Schiff base; Glycolysis; Multigene family.
 FT INIT MET 0
 FT BINDING 55 55 BY SIMILARITY.
 FT BINDING 146 146 C-1-PHOSPHATE GROUP OF THE SUBSTRATE.
 FT BINDING 229 229 SCHIFF-BASE WITH DIHYDROXYACETONE-P.
 FT ACT_SITE 363 363 ESSENTIAL FOR ENHANCED ACTIVITY OF THE ENZYME TOWARD FRUCTOSE 1,6-BISPHOSPHATE AS COMPARED WITH FRUCTOSE 1-PHOSPHATE.

CONFLICT 80 80 H -> Y (IN REF. 2; AAH34173).
 FT CONFLICT 154 155 AD -> DH (IN REF. 1).
 FT SEQUENCE 363 AA; 39375 MW; 891996AABE94B025 CRC64;

Query Match 75.7%; Score 28; DB 1; Length 363;
 Best Local Similarity 66.7%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 DB 349 GAAATGSLF 357

Search completed: January 29, 2004, 14:53:09

Fri Jan 30 09:06:42 2004

us-10-032-950-3.rsp

Page 11

Job time : 6.82353 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:46:19 ; Search time 26.8676 Seconds
(without alignments)
86.441 Million cell updates/sec

Title: US-10-032-950-3

Perfect score: 37

Sequence: 1 GVAAXSLF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	91.9	377	10 Q8L676	Q8L676 oryza sativ
2	33	89.2	1253	5 Q24982	Q24982 giardia lam
3	32	86.5	385	5 Q88RY6	Q88RY6 encphalito
4	31	83.8	163	16 Q9CCE8	Q9CCE8 mycobacteri
5	31	83.8	167	16 Q53195	Q53195 mycobacteri
6	31	83.8	169	2 Q9X7B5	Q9X7B5 mycobacteri
7	31	83.8	255	10 Q9LE76	Q9LE76 oryza sativ
8	30	81.1	226	2 Q56969	Q56969 yersinia pe
9	30	81.1	267	16 Q916F3	Q916F3 pseudomonas
10	30	81.1	305	16 Q8UF26	Q8UF26 agrobacteri
11	30	81.1	565	6 Q77540	Q77540 orycolagus
12	29	78.4	288	5 Q85V43	Q85V43 encphalito
13	29	78.4	335	16 Q98G45	Q98G45 rhizobium 1
14	29	78.4	349	5 Q9N6S1	Q9N6S1 caenorhabdi
15	29	78.4	358	10 Q9XEC5	Q9XEC5 arabidopsis
16	29	78.4	403	17 Q9HKZ8	Q9HKZ8 thermoplas

17	29	78.4	418	17 Q97389	Q97389 thermoplas
18	29	78.4	536	5 Q8U51	Q8U51 encphalito
19	29	78.4	626	5 Q76144	Q76144 neospora ca
20	29	78.4	654	10 Q9AVL8	Q9AVL8 oryza sativ
21	29	78.4	1668	17 Q27011	Q27011 methanobact
22	28	75.7	42	13 Q92007	Q92007 gallus gall
23	28	75.7	47	16 Q33586	Q33586 staphylococ
24	28	75.7	169	12 Q917R7	Q917R7 hepatitis c
25	28	75.7	173	8 Q8HLA7	Q8HLA7 hypopitichus
26	28	75.7	249	3 Q9F911	Q9F911 comanonas s
27	28	75.7	258	3 Q93871	Q93871 verticilliu
28	28	75.7	260	4 Q96B15	Q96B15 homo sapien
29	28	75.7	263	4 Q9BWD9	Q9BWD9 homo sapien
30	28	75.7	267	2 Q9ROP1	Q9ROP1 vibrio chol
31	28	75.7	274	2 Q8KX48	Q8KX48 cennococcus
32	28	75.7	298	16 Q8DQ32	Q8DQ32 synchococ
33	28	75.7	307	2 Q9URE6	Q9URE6 xanthomonas
34	28	75.7	327	16 Q9BDQ4	Q9BDQ4 rhizobium 1
35	28	75.7	331	10 Q41628	Q41628 triticum tu
36	28	75.7	331	10 Q06978	Q06978 zea mays (m
37	28	75.7	337	2 Q48593	Q48593 lactococcus
38	28	75.7	338	2 Q56925	Q56925 yersinia en
39	28	75.7	339	16 Q9ZC04	Q9ZC04 yersinia pe
40	28	75.7	348	8 Q47404	Q47404 poecilia m
41	28	75.7	348	8 Q47405	Q47405 poecilia ar
42	28	75.7	364	11 Q91Y97	Q91Y97 mus musculu
43	28	75.7	364	11 Q8K034	Q8K034 mus musculu
44	28	75.7	364	11 Q9C8C1	Q9C8C1 mus musculu
45	28	75.7	364	11 Q91W73	Q91W73 mus musculu

ALIGNMENTS

RESULT 1

Q8L676 ID Q8L676 PRELIMINARY; PRT; 377 AA.

AC Q8L676; 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypochemical protein.
 GN OSUNBA0011109.21.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OC NCBI_TaxID=39947;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,
 RA Overton II L.V., Bera J.J., Teltrin T., Krol M.I., Jarrah B.B.,
 RA Jin S.S., Koo H., Ziemann V., Hsiao J., Blunt S., Vanaken S.S.,
 RA Uterback T.T., Feldblyum T.V., Yang Q.Q., Haas B.J., Suh B.B.,
 RA Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.,
 RT "Oryza sativa chromosome 10 BAC OSUNBA0011109 genomic sequence."
 RL Submitted (May-2002) to the EMBL/GenBank/DBS databases.
 DR EMBL; AC092388; AAM22720.1; -
 DR Gramene; Q8L676; -
 DR InterPro; IPR004843; M-Pestrase.
 DR Pfam; PF00149; Metallophos; 1.
 KW Hypochemical Protein.
 SQ SEQUENCE 377 AA; 40656 MW; C18C31695DB5B459 CRC64;

Query Match 91.9%; Score 34; DB 10; Length 377;

Best Local Similarity 77.8%; Pred. No. 13;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 DB 239 GVAAXSLF 247

```

RESULT 2
Q24982 PRELIMINARY; PRT; 1253 AA.
AC Q24982;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Pyruvate:ferredoxin oxidoreductase.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RA Upcroft P., Healey A., Upcroft J., Thomson S.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94170792; PubMed=8125101;
RA Thomson S.M., Hanson G.R., Upcroft J.A., Upcroft P.;
RT "A purified ferredoxin from Giardia duodenalis.";
RU Eur J Biochem. 220:439-446(1994).
DR EMBL, L27221; AAA74894.1; -
DR InterPro; IPR001450; 4Fe4S_Ferredoxin.
DR InterPro; IPR002869; POR_N_D2.
DR InterPro; IPR002880; POR_N.
DR Pfam; PF00037; Fe4S_2.
DR Pfam; PF01558; POR_1.
DR Pfam; PF01855; POR_N_1.
DR PROSITE; PS00198; 4Fe4S_FERREDOXIN; 2.
DR PROSITE; PS00667; CPASASE_2; 1.
DR 4Fe-4S; Iron; Iron-sulfur.
SQ SEQUENCE 1253 AA; 138928 MW; F3B74820DEA8B21 CRC64;

Query Match
Best Local Similarity 89.2%; Score 33; DB 5; Length 1253;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

CY 1 GVAAXSSLF 9
DB 906 GAAACSLF 914

RESULT 3
Q8SRV6 PRELIMINARY; PRT; 385 AA.
AC Q8SRV6;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Putative AMINOACID transporter.
GN ECTO5_0580.
OS Encephalitozoon cuniculi.
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=GB-M1;
RC Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornilliot E., Metenier G., Thamarat F.,
RA Prensili G., Barbe V., Peyrethallade E., Brothier P., Winkler P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi.";
RL Nature 414:450-453(2001).
DR EMBL, AL590445; CAD26577.1; -

```

```

DR InterPro; IPR002422; AA/rel_permease2.
DR Pfam; PF01490; Aa_trans; 1.
DR Hypothetical protein.
SQ SEQUENCE 385 AA; 43606 MW; 0EBB01706C82ECAB CRC64;

Query Match
Best Local Similarity 86.5%; Score 32; DB 5; Length 385;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

CY 1 GVAAXSSLF 9
DB 374 GVAAPASLF 382

RESULT 4
Q9CC58 PRELIMINARY; PRT; 163 AA.
AC Q9CC58;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Hypothetical protein ML1255.
GN ML1255.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=TN;
RX MEDLINE=2118732; PubMed=11234002;
RA Cole S.T., Sigmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basam D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moulé S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream K.A., Rutherford K.W.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrall B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
DR EMBL, AL583921; CAC31636.1; -
DR Lepidoma; ML1255; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 163 AA; 16842 MW; F26A1A2C04E3FD2E CRC64;

Query Match
Best Local Similarity 83.8%; Score 31; DB 16; Length 163;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 GVAAXSSLF 9
DB 132 GVAASASLF 140

RESULT 5
Q53195 PRELIMINARY; PRT; 167 AA.
AC Q53195;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein RV2468C.
GN RV2468C OR MT008.24C OR MT2543.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=H37Rv;
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;

```

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala P.,
 RA Baddock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares R.,
 RA Stinton J.E., Taylor K., Whitehead S., Barrall B.G.,
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.",
 RL Nature 393:537-544(1998).
 RN [2].
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolony J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.",
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL021246; CA16045.1; -;
 DR EMBL; AB007091; AAK46843.1; -;
 DR TIGR; MT2543; -;
 DR TubercuList; RV2468C; -;
 KM Hypothesized protein; Complete proteome.
 SQ SEQUENCE 167 AA; 1288 MW; F600B6FC5A5E4B4 CRC64;

Query Match 83.8%; Score 31; DB 16; Length 167;
 Best Local Similarity 77.8%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 GVAAXSSLF 9
 Db 136 GVAAXSSAF 144

RESULT 6
 ID Q9X7B5 PRELIMINARY; PRT; 169 AA.
 AC Q9X7B5;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothesized 17.4 kDa protein.
 GN MLCB1610.16.
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1].
 RP SEQUENCE FROM N.A.
 RA Seeger K.J., Harris D.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2].
 RP SEQUENCE FROM N.A.
 RA James K.D., Parkhill J., Barrall B.G., Rajandream M.A.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3].
 RP SEQUENCE FROM N.A.
 RC MEDLINE=93188700; Pubmed=8446027;
 RA Eiglmeier K., Honore N., Woods S.A., Caudron B., Cole S.T.,
 RT "Use of an ordered cosmid library to deduce the genomic organization
 RT of Mycobacterium leprae.",
 RL Mol. Microbiol. 7:197-206(1993).
 DR EMBL; AL049913; CAB43162.1; -;
 KM Hypothesized protein.
 SQ SEQUENCE 169 AA; 17360 MW; 2FB50C9B845D0B9 CRC64;

Query Match 83.8%; Score 31; DB 2; Length 169;
 Best Local Similarity 77.8%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 GVAAXSSLF 9
 Db 138 GVAAXSSAF 146

RESULT 7
 ID Q9LE76 PRELIMINARY; PRT; 255 AA.
 AC Q9LE76;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Hypothesized protein.
 OS Oryza sativa (rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1].
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone: P0693B08.",
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP001081; BAA90367.1; -;
 DR EMBL; AP001073; BAA89586.1; -;
 DR Gramene; Q9LE76; -;
 SQ SEQUENCE 255 AA; 26887 MW; 0CA5C39B9987549 CRC64;

Query Match 83.8%; Score 31; DB 10; Length 255;
 Best Local Similarity 77.8%; Pred. No. 43;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 GVAAXSSLF 9
 Db 82 GVAAXSSAF 90

RESULT 8
 ID Q56969 PRELIMINARY; PRT; 236 AA.
 AC Q56969;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE LcrK protein (Fragment).
 GN LCRKA.
 OS Yersinia pestis.
 OG Plasmid lcr.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN [1].
 RP SEQUENCE OF 142-236 FROM N.A.
 RC STRAIN=358;
 RA Filippov A.A., Oleinikov P.N., Motin V.I., Protzenko O.A.,
 RT "Sequencing of two Yersinia pestis IS elements, IS285 and IS100",
 RL Contrib. Microbiol. Immunol. 13:0-0(1995).
 RN [2].
 RP SEQUENCE FROM N.A.
 RC STRAIN=358;
 RA Filippov A.A.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.

```

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=358;
RX MEDLINE=92250432; PubMed=1577700;
RA Rimpilaenen M., Forsberg A., Wolf-Matz H.;
RT "A novel protein, LcrG, involved in the low-calcium response of
  Yersinia pseudotuberculosis shows extensive homology to YopH.";
RL J. Bacteriol. 174:3355-3365(1992).
DR EMBL: X78303; CAAS5109.1; -
DR InterPro: IPR003282; SecIiOMPX.
DR InterPro: IPR006182; YscJ FlIF.
DR Pfam: PF01514; YscJ FlIF.
DR PRINTS: PR01338; TYPE3OMKPROT.
KM Plasmid.
FT NON_TER
FT CONFLICT
SQ SEQUENCE 236 AA; 26108 MM; 3E12AFB692E1373F CRC64;

Query Match 81.1%; Score 30; DB 2; Length 236;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVAXSSLP 9
Db 135 GVAXSASVF 143

RESULT 9
Q916F3 PRELIMINARY; PRT; 267 AA.
ID Q916F3;
AC Q916F3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein PA0340.
GN PA0340.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Steyer C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino B., Westbrook-Wadman S., Yan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reiter J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
  opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL: AB004472; AAG03729.1; -
DR InterPro: IPR002781; DUF81.
DR Pfam: PF01925; DUF81.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 267 AA; 27809 MM; DA435AF6A616EB01 CRC64;

Query Match 81.1%; Score 30; DB 16; Length 267;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAXSSLP 9
Db 219 GIATSSLP 227

RESULT 10
Q8UFZ6 PRELIMINARY; PRT; 305 AA.
AC Q8UFZ6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)

```

```

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein Atu1248.
GN ATU1248 OR AGR_C 2304.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D.S.,
RA Chapman P., Clendenning J., Deatherage G., Gillet M., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphitumachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
  C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Homiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughy D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
  Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL: AB009087; AL442257.1; -
DR EMBL: AB008052; AA87045.1; -
DR InterPro: IPR01525; CS DNA_meth.
DR InterPro: IPR000620; DUF6.
DR Pfam: PF00892; DUF6.
DR PROSITE: PS00095; CS_MTASE_2; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 305 AA; 32130 MM; 94EB9C8C22A2PDDF CRC64;

Query Match 81.1%; Score 30; DB 16; Length 305;
Best Local Similarity 87.5%; Pred. No. 89;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VAXSSLP 9
Db 127 VAAATSLP 134

RESULT 11
ID 077540 PRELIMINARY; PRT; 565 AA.
AC 077540;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Liver carboxylesterase (EC 3.1.1.1).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver.
RX MEDLINE=98297515; PubMed=9635592;
RA Potter P.M., Pavlik C.A., Norton C.L., Naeve C.W., Danks M.K.;
RT "Isolation and partial characterization of a cDNA encoding a rabbit
  liver carboxylesterase that activates the prodrug irinotecan (CPT-

```

```

RL 11) ".
CC Cancer Res. 58:2646-2651(1998).
CC -1- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
DR EMBL; AF036930; AAC39258.1; -.
DR HSSP; P21836; IMAH.
DR InterPro; IPR002018; Carboxylesterase.
DR InterPro; IPR000379; Ser esters_alte.
DR Pfam; PF00135; Coesterase; 1.
DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
DR HydroLase.
SQ SEQUENCE 565 AA; 62291 MW; 0ACD61400CC81D2F CRC64;

Query Match
Best Local Similarity 81.4%; Score 30; DB 6; Length 565;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVAAXSLF 9
DB 248 GVALLSLF 256

RESULT 12
Q8SVA3 PRELIMINARY; PRT; 298 AA.
AC Q8SVA3;
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical protein EC006_0970.
GN EC006_0970.
OS Eucephalitozoon cuniculi.
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Eucephalitozoon.
OX NCBI_TaxID=6035;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=GB-M1;
RC MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornillot E., Mecerat G., Thomarat F.,
RA Prensier G., Barbe V., Peyretailade E., Broctier P., Wincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Eucephalitozoon cuniculi.";
RL Nature 414:450-453(2001).
DR EMBL; AL590446; CAD25457.1; -.
KW Hypothetical protein.
SQ SEQUENCE 298 AA; 32992 MW; 82ABE7B530F4395 CRC64;

Query Match
Best Local Similarity 78.4%; Score 29; DB 5; Length 298;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVAAXSLF 9
DB 134 GVAATNLF 142

RESULT 13
Q98G45 PRELIMINARY; PRT; 335 AA.
AC Q98G45;
DT 01-OCT-2001 (TREMblrel. 18, Created)
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Iron-binding periplasmic protein, SfuA protein precursor.
GN M2J3495.
OS Rhizobium loti (Mesorhizobium loti).

```

```

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=MAFF30309;
RC MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asanizu E., Kato T., Saito S.,
RA Watanabe A., Iiesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003002; BAB50371.1; -.
DR InterPro; IPR006059; SBP_bac_1.
DR Pfam; PF01547; SBP_bac_1; 1.
DR Complete proteome.
SQ SEQUENCE 335 AA; 35101 MW; 722FFFA124A6F1C0 CRC64;

Query Match
Best Local Similarity 78.4%; Score 29; DB 16; Length 335;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GVAAXSLF 9
DB 123 GVAARSTVF 131

RESULT 14
Q9N4S1 PRELIMINARY; PRT; 349 AA.
AC Q9N4S1;
ID Q9N4S1;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Hypothetical 37.7 kDa protein.
GN Y53GB.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC Latreille P.;
RT "The sequence of C. elegans cosmid Y53GB.4.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC Waterston R.;
RT "Direct Submision.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006804; AAF60757.1; -.
DR WormPep; Y53GB.4; CE25430.
KW Hypothetical protein.
SQ SEQUENCE 349 AA; 37676 MW; C0E76887720CA0AB CRC64;

Query Match
Best Local Similarity 78.4%; Score 29; DB 5; Length 349;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVAAXSLF 9

```

Db 61 GLVSSSLF 69

RESULT 15

Q9XEC5 PRELIMINARY; PRT; 398 AA.

ID Q9XEC5

DT 01-NOV-1999 (TRENBLrel. 12, Created)

DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)

DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)

DE Hypothetical 44.0 kDa protein.

GN T26N6.8 OR A74G04480.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eucots II; Brassicales; Brassicaceae; Arabidopsis.

NCBI_TaxID=3702;

[1] SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RT Zhong J., Ma P., Parnell L.D., Chen C.N., Chen E.Y.,

RT "Genomic sequence of Arabidopsis thaliana BAC T26N6, chromosome IV,

RT 19.3 cM."

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

[2] SEQUENCE FROM N.A.

RP Zhong J., Ma P., Parnell L.D., Chen C.N., Chen E.Y.,

RP Lemcke K., Mayer K.F.X.,

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[3] SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF076243; AAD29750.1; -

DR EMBL; AL161500; CAB77916.1; -

KW Hypothetical protein.

SQ SEQUENCE 398 AA; 43961 MW; 007BE2298E94AD7 CRC64;

Query Match 78.4%; Score 29; DB 10; Length 398;

Best Local Similarity 66.7%; Pred. No. 2e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLVSSSLF 9

Db 289 GLVSSSLF 297

Search completed: January 29, 2004, 14:56:46
Job time : 27.8676 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:52:19 : Search time 11.1176 Seconds
(without alignments)
34.252 Million cell updates/sec

Title: US-10-032-950-3

Perfect score: 37

Sequence: 1 GVAXXSLF 9

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 4231058 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued Patents AA:*
2: /cgn2_6/prodata/1/iaa/5A-COMB.pep:*
3: /cgn2_6/prodata/1/iaa/5B-COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6A-COMB.pep:*
5: /cgn2_6/prodata/1/iaa/6B-COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfillseq.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	94.6	9	4 US-09-339-511-3	Sequence 3, Appli
2	30	81.1	332	4 US-09-252-991A-31791	Sequence 31791, A
3	30	81.1	347	4 US-09-107-532A-5861	Sequence 5861, Ap
4	30	81.1	539	4 US-09-264-737-1	Sequence 1, Appli
5	30	81.1	566	3 US-09-264-737-2	Sequence 2, Appli
6	29	78.4	9	4 US-09-339-511-1	Sequence 1, Appli
7	28	75.7	9	4 US-08-861-476C-3	Sequence 3, Appli
8	28	75.7	20	1 US-07-988-430-59	Sequence 59, Appli
9	28	75.7	20	1 US-08-425-336-57	Sequence 57, Appli
10	28	75.7	20	1 US-08-488-118-57	Sequence 57, Appli
11	28	75.7	20	1 US-08-477-684B-57	Sequence 57, Appli
12	28	75.7	20	1 US-08-646-560-57	Sequence 57, Appli
13	28	75.7	20	3 US-08-839-765-57	Sequence 57, Appli
14	28	75.7	20	3 US-09-136-389-57	Sequence 57, Appli
15	28	75.7	20	4 US-09-610-838-57	Sequence 57, Appli
16	28	75.7	20	5 PCT-US92-09487-59	Sequence 59, Appli
17	28	75.7	47	4 US-08-661-476C-6	Sequence 59, Appli
18	27	73.0	448	4 US-09-328-352-5807	Sequence 5807, Ap
19	27	73.0	530	4 US-09-252-991A-31791	Sequence 28311, A
20	27	73.0	791	1 US-08-394-880B-2	Sequence 2, Appli
21	26	70.3	9	4 US-09-339-511-5	Sequence 5, Appli
22	26	70.3	9	4 US-09-339-511-6	Sequence 6, Appli
23	26	70.3	258	4 US-09-556-916-30	Sequence 30, Appli
24	26	70.3	258	4 US-09-556-916-32	Sequence 32, Appli
25	26	70.3	365	4 US-09-252-991A-21955	Sequence 21955, A
26	26	70.3	439	4 US-09-252-991A-28440	Sequence 28440, A
27	26	70.3	547	1 US-08-083-948-8	Sequence 8, Appli

28	26	70.3	547	1	US-08-393-785-8	Sequence 8, Appli
29	26	70.3	547	1	US-08-473-694-8	Sequence 8, Appli
30	26	70.3	547	1	US-08-712-057-8	Sequence 8, Appli
31	26	70.3	547	1	US-09-347-878-36	Sequence 36, Appli
32	26	70.3	605	4	US-09-252-991A-32874	Sequence 32874, A
33	26	70.3	672	4	US-09-556-916-26	Sequence 26, Appli
34	26	70.3	672	4	US-09-556-916-28	Sequence 28, Appli
35	26	70.3	675	4	US-09-252-991A-39161	Sequence 29161, A
36	26	70.3	909	4	US-09-252-991A-30503	Sequence 30503, A
37	26	70.3	919	4	US-09-437-054A-17	Sequence 17, Appli
38	26	67.6	9	4	US-09-339-511-2	Sequence 2, Appli
39	25	67.6	108	4	US-09-199-637A-227	Sequence 227, App
40	25	67.6	110	4	US-09-107-532A-4192	Sequence 4192, App
41	25	67.6	114	4	US-09-615-192A-110	Sequence 110, App
42	25	67.6	127	4	US-09-198-452A-1034	Sequence 1034, App
43	25	67.6	138	3	US-09-422-869-16	Sequence 16, Appli
44	25	67.6	139	3	US-09-422-869-14	Sequence 14, Appli
45	25	67.6	141	4	US-09-328-352-6288	Sequence 6288, Ap

ALIGNMENTS

```

RESULT 1
US-09-339-511-3
; Sequence 3, Application US/09339511
; Patent No. 6337385
; GENERAL INFORMATION:
; APPLICANT: Mair, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. 6337385ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; PUB REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/09/339, 511
; CURRENT FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: 60/090, 402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-3
Query Match          94.6% Score 35; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. NO. 2.5e+05;
Matches          9; Conservative          0; Mismatches          0; Indels          0; Gaps          0;
QY          1 GVAXXSLF 9
Db          1 GVAXXSLF 9
RESULT 2
US-09-252-991A-31791
; Sequence 31791, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252, 991A

```

;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 31791
;; LENGTH: 332
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31791

Query Match 81.1%; Score 30; DB 4; Length 332;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLSLF 9
Db 284 GAAATSMFLF 292

RESULT 3
US-09-107-532A-5861
; Sequence 5861, Application US/09107532A
; Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: LYNN A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Arinello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 5861:

SEQUENCE CHARACTERISTICS:

LENGTH: 347 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...347

SEQUENCE DESCRIPTION: SEQ ID NO: 5861:

US-09-107-532A-5861

Query Match 81.1%; Score 30; DB 4; Length 347;

Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLSLF 9
Db 87 GAAAGSIF 95

RESULT 4
US-09-264-737-1
; Sequence 1, Application US/09264737A
; Patent No. 6107549

GENERAL INFORMATION:

APPLICANT: Feng, Paul C.C.

TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via

FILE REFERENCE: 38-21(10551) RLE3 Pyridine Tolerance

CURRENT FILING DATE: 1999-03-09

EARLIER APPLICATION NUMBER: 60/077,377

EARLIER FILING DATE: 1998-03-10

NUMBER OF SEQ ID NOS: 11

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1

LENGTH: 539

TYPE: PRT

ORGANISM: Rabbit

US-09-264-737-1

Query Match 81.1%; Score 30; DB 3; Length 539;
Best Local Similarity 77.8%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLSLF 9
Db 222 GVALLSLF 230

RESULT 5

US-09-264-737-2

Sequence 2, Application US/09264737A

Patent No. 6107549

GENERAL INFORMATION:

APPLICANT: Feng, Paul C.C.

TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via

FILE REFERENCE: 38-21(10551) RLE3 Pyridine Tolerance

CURRENT FILING DATE: 1999-03-09

EARLIER APPLICATION NUMBER: 60/077,377

EARLIER FILING DATE: 1998-03-10

NUMBER OF SEQ ID NOS: 11

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 566

TYPE: PRT

ORGANISM: Rabbit

US-09-264-737-2

Query Match 81.1%; Score 30; DB 3; Length 566;
Best Local Similarity 77.8%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLSLF 9
Db 249 GVALLSLF 257

RESULT 6

US-09-339-511-1
; Sequence 1, Application US/09339511

Patent No. 6337385
 GENERAL INFORMATION:
 APPLICANT: Malt, Tom
 APPLICANT: Mayville, Patricia
 APPLICANT: No. 6337385ick, Richard P.
 APPLICANT: Beavis, Ronald
 APPLICANT: Ji, Guangyong
 TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
 TITLE OF INVENTION: INTERFERENCE
 FILE REFERENCE: 600-1-231N
 CURRENT APPLICATION NUMBER: US/09/339,511
 CURRENT FILING DATE: 1999-06-24
 PRIOR APPLICATION NUMBER: 60/090,402
 PRIOR FILING DATE: 1998-06-24
 NUMBER OF SEQ. ID NOS.: 8
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1
 LENGTH: 9
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 OTHER INFORMATION: peptide
 NAME/KEY: VARIANT
 LOCATION: (5)
 OTHER INFORMATION: Xaa represents any amino acid at this position.
 US-09-339-511-1

Query Match 78.4%; Score 29; DB 4; Length 9;
 Best Local Similarity 88.9%; Pred. No. 2.5e+05;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVAAXSLSF 9
 DB 1 GVAAXSLSF 9

RESULT 7
 US-08-861-476C-3
 Sequence 3, Application US/08861476C
 Patent No. 6447786
 GENERAL INFORMATION:
 APPLICANT: New York University Medical Center
 TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
 FILE REFERENCE: 63753/7
 CURRENT APPLICATION NUMBER: US/08/861,476C
 CURRENT FILING DATE: 1997-05-22
 NUMBER OF SEQ. ID NOS.: 8
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 3
 LENGTH: 9
 TYPE: PRT
 ORGANISM: Staphylococcus aureus
 US-08-861-476C-3

Query Match 75.7%; Score 28; DB 4; Length 9;
 Best Local Similarity 77.8%; Pred. No. 2.5e+05;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLSF 9
 DB 1 GVAAXSLSF 9

RESULT 8
 US-07-988-430-59
 Sequence 59, Application US/07988430
 Patent No. 5416202
 GENERAL INFORMATION:
 APPLICANT: Bernhard, Susan L.
 APPLICANT: Better, Marc D.
 APPLICANT: Carroll, Stephen F.
 APPLICANT: Lane, Julie A.

APPLICANT: Lei, Shau-Ping
 TITLE OF INVENTION: Materials Comprising and Methods of
 TITLE OF INVENTION: Preparation and Use for Ribosome-Inactivating Proteins
 NUMBER OF SEQUENCES: 101
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ADDRESSEE: Bicknell
 STREET: Two First National Plaza, 20 South Clark
 STREET: Street
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60603
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/988,430
 FILING DATE: 19921209
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/901,707
 FILING DATE: 19-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/787,567
 FILING DATE: 04-NOV-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5416202and, Greta E.
 REGISTRATION NUMBER: 35302
 REFERENCE/DOCKET NUMBER: 31133
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 346-5750
 TELEFAX: (312) 984-9740
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 59:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-07-988-430-59

Query Match 75.7%; Score 28; DB 1; Length 20;
 Best Local Similarity 66.7%; Pred. No. 2.7;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVAAXSLSF 9
 DB 6 GAAASESLF 14

RESULT 9
 US-08-425-336-57
 Sequence 57, Application US/08425336
 Patent No. 5621083
 GENERAL INFORMATION:
 APPLICANT: Better, Marc D.
 APPLICANT: Carroll, Stephen F.
 APPLICANT: Studinka, Gary M.
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 TITLE OF INVENTION: Proteins
 NUMBER OF SEQUENCES: 140
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/425,336
 FILING DATE: 18-APR-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/064,691
 FILING DATE: 12-MAY-1993
 APPLICATION NUMBER: US 07/901,707
 FILING DATE: 19-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/787,567
 FILING DATE: 04-NOV-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Meyers, Thomas C.
 REGISTRATION NUMBER: P-36,989
 REFERENCE/DOCKET NUMBER: 31394
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 57:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-425-336-57

Query Match 75.7%; Score 28; DB 1; Length 20;
 Best Local Similarity 66.7%; Pred. No. 2.7;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 |||||
 DB 6 GAASESLF 14

RESULT 10
 US-08-488-113B-57
 ; Sequence 57, Application US/08488113B
 ; Patent No. 5744580
 ; GENERAL INFORMATION:
 ; APPLICANT: Better, Marc D.
 ; APPLICANT: Carroll, Stephen F.
 ; APPLICANT: Studnika, Gary M.
 ; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 ; NUMBER OF SEQUENCES: 169
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
 ; STREET: 500 West Madison Street, 34th floor
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60661
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/488,113B
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/425,336
 ; FILING DATE: 18-APR-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/064,691
 ; FILING DATE: 12-MAY-1993

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/988,430
 FILING DATE: 09-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/901,707
 FILING DATE: 19-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/787,567
 FILING DATE: 04-NOV-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: McNicholas, Janet M.
 REGISTRATION NUMBER: 32,918
 REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/707-8889
 TELEFAX: 312/707-9155
 TELEX: 650 388-1248
 INFORMATION FOR SEQ ID NO: 57:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-488-113B-57

Query Match 75.7%; Score 28; DB 1; Length 20;
 Best Local Similarity 66.7%; Pred. No. 2.7;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 |||||
 DB 6 GAASESLF 14

RESULT 11
 US-08-477-484B-57
 ; Sequence 57, Application US/08477484B
 ; Patent No. 5756699
 ; GENERAL INFORMATION:
 ; APPLICANT: Better, Marc D.
 ; APPLICANT: Carroll, Stephen F.
 ; APPLICANT: Studnika, Gary M.
 ; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating.
 ; NUMBER OF SEQUENCES: 169
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
 ; STREET: 500 West Madison Street, 34th floor
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60661
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/477,484B
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/425,336
 ; FILING DATE: 18-APR-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/064,691
 ; FILING DATE: 12-MAY-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/988,430
 ; FILING DATE: 09-DEC-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/901,707
 ; FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/787,567
 FILING DATE: 04-NOV-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: McNicholas, Janet M.
 REGISTRATION NUMBER: 32,918
 REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/707-8889
 TELEFAX: 312/707-9155
 TELEX: 650 388-1248
 INFORMATION FOR SEQ ID NO: 57:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-477-484B-57

Query Match 75.7%; Score 28; DB 1; Length 20;
 Best Local Similarity 66.7%; Pred. No. 2.7;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 Db 6 GAAXSESLF 14

RESULT 12
 US-08-646-360-57
 Sequence 57, Application US/08646360
 Patent No. 5837491
 GENERAL INFORMATION:
 APPLICANT: Better, Marc D.
 APPLICANT: Carroll, Stephen F.
 APPLICANT: Studnika, Gary M.
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 TITLE OF INVENTION: Proteins
 NUMBER OF SEQUENCES: 173
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McAndrews, Held & Malloy, Ltd.
 STREET: 500 West Madison Street, 34th floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60661
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/646,360
 FILING DATE: 13-MAY-1996
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/05348
 FILING DATE: 12-MAY-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/064,691
 FILING DATE: 12-MAY-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/988,430
 FILING DATE: 09-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/901,707
 FILING DATE: 19-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/787,567
 FILING DATE: 04-NOV-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: McNicholas, Janet M.
 REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.P4
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/707-8889
 TELEFAX: 312/707-9155
 TELEX: 650 388-1248
 INFORMATION FOR SEQ ID NO: 57:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-646-360-57

Query Match 75.7%; Score 28; DB 2; Length 20;
 Best Local Similarity 66.7%; Pred. No. 2.7;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
 Db 6 GAAXSESLF 14

RESULT 13
 US-08-839-765-57
 Sequence 57, Application US/08839765
 Patent No. 614631
 GENERAL INFORMATION:
 APPLICANT: Better, Marc D.
 APPLICANT: Carroll, Stephen F.
 APPLICANT: Studnika, Gary M.
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 TITLE OF INVENTION: Proteins
 NUMBER OF SEQUENCES: 169
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McAndrews, Held & Malloy, Ltd.
 STREET: 500 West Madison Street, 34th floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60661
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/839,765
 FILING DATE: 15-APR-1997
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/425,336
 FILING DATE: 18-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/064,691
 FILING DATE: 12-MAY-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/988,430
 FILING DATE: 09-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/901,707
 FILING DATE: 19-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/787,567
 FILING DATE: 04-NOV-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: McNicholas, Janet M.
 REGISTRATION NUMBER: 32,918
 REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/707-8889
 TELEFAX: 312/707-9155
 TELEX: 650 388-1248
 INFORMATION FOR SEQ ID NO: 57:

SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-839-765-57

Query Match 75.7%; Score 28; DB 3; Length 20;
Best Local Similarity 66.7%; Pred. No. 2.7;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
| | | | |
Db 6 GAASESLF 14

RESULT 14
US-09-136-389-57
Sequence 57, Application US/09136389
Patent No. 6146850
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
APPLICANT: Carroll, Stephen F.
APPLICANT: Studnika, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
NUMBER OF INVENTION: Proteins
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,389
CLASSIFICATION:
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-09-136-389-57

Query Match 75.7%; Score 28; DB 3; Length 20;
Best Local Similarity 66.7%; Pred. No. 2.7;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
| | | | |
Db 6 GAASESLF 14

RESULT 15
US-09-610-838-57
Sequence 57, Application US/09610838
Patent No. 6376217
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
APPLICANT: Carroll, Stephen F.
APPLICANT: Studnika, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
NUMBER OF INVENTION: Proteins
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/610,838
CLASSIFICATION:
FILING DATE: 06-JUL-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/136,389
FILING DATE: 18-AUG-1998
APPLICATION NUMBER: 08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-610-838-57

Query Match 75.7%; Score 28; DB 4; Length 20;
Best Local Similarity 66.7%; Pred. No. 2.7;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GVAAXSSLF 9
| | | | |
Db 6 GAASESLF 14

Search completed: January 29, 2004, 15:03:21
Job time : 12.1176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:47:39 ; Search time 25.5441 Seconds
(without alignments)
73.223 Million cell updates/sec

Title: US-10-032-950-3

Perfect score: 37

Sequence: 1 GVAAXSLF 9

Scoring table: BLOSUM62

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database:

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PTCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PTCT_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	94.6	9	13 US-10-032-950-3	Sequence 3, Appli
2	32	86.5	391	15 US-10-156-761-9222	Sequence 9222, Ap
3	31	83.8	163	16 US-10-080-170-155	Sequence 155, Ap
4	31	83.8	167	16 US-10-080-170-538	Sequence 538, Ap
5	30	81.1	565	12 US-10-267-786-2	Sequence 2, Appli
6	29	78.4	9	13 US-10-032-950-1	Sequence 1, Appli
7	29	78.4	160	10 US-09-966-546-10	Sequence 10, Appli
8	29	78.4	160	10 US-09-966-545-10	Sequence 10, Appli
9	29	78.4	160	10 US-09-966-212-10	Sequence 10, Appli
10	29	78.4	160	16 US-10-189-940-10	Sequence 10, Appli
11	29	78.4	161	10 US-09-966-546-12	Sequence 12, Appli
12	29	78.4	161	10 US-09-966-545-12	Sequence 12, Appli
13	29	78.4	161	11 US-09-965-212-12	Sequence 12, Appli
14	29	78.4	161	11 US-10-189-940-12	Sequence 12, Appli
15	29	78.4	1668	12 US-10-369-493-1181	Sequence 1181, Ap

16	28	75.7	9	15 US-10-201-444-3	Sequence 3, Appli
17	28	75.7	20	12 US-10-127-890-57	Sequence 57, Appli
18	28	75.7	31	9 US-09-925-301-1462	Sequence 1462, Ap
19	28	75.7	15	15 US-10-201-444-6	Sequence 6, Appli
20	28	75.7	47	15 US-10-264-049-3838	Sequence 3838, Ap
21	28	75.7	364	11 US-09-919-039-343	Sequence 343, Ap
22	28	75.7	364	12 US-10-205-719-67	Sequence 67, Appli
23	28	75.7	489	15 US-10-156-761-12438	Sequence 12438, A
24	28	75.7	800	12 US-10-369-493-14064	Sequence 14064, A
25	27	73.0	163	12 US-10-094-749-2459	Sequence 2459, Ap
26	27	73.0	238	12 US-10-369-493-10775	Sequence 10775, A
27	27	73.0	295	12 US-10-369-493-19090	Sequence 19090, A
28	27	73.0	302	12 US-10-369-493-12198	Sequence 12198, A
29	27	73.0	327	10 US-09-738-626-5251	Sequence 5251, Ap
30	27	73.0	337	12 US-10-369-493-9082	Sequence 9082, Ap
31	27	73.0	475	15 US-10-156-761-13178	Sequence 13178, A
32	27	73.0	478	15 US-10-156-761-13018	Sequence 13018, A
33	27	73.0	522	15 US-10-156-761-9801	Sequence 9801, Ap
34	27	73.0	558	15 US-10-156-761-10602	Sequence 10602, A
35	27	73.0	604	12 US-10-369-493-14086	Sequence 4086, Ap
36	27	73.0	607	12 US-09-949-029-6	Sequence 6, Appli
37	27	73.0	935	12 US-10-128-202-7	Sequence 7, Appli
38	27	73.0	945	12 US-10-128-202-8	Sequence 8, Appli
39	27	73.0	946	12 US-10-085-198-128	Sequence 128, Ap
40	27	73.0	959	12 US-10-128-202-2	Sequence 2, Appli
41	26	70.3	9	13 US-10-032-950-5	Sequence 5, Appli
42	26	70.3	9	13 US-10-032-950-6	Sequence 44910, A
43	26	70.3	50	9 US-09-864-761-44910	Sequence 1919, Ap
44	26	70.3	125	12 US-10-264-237-1919	Sequence 8741, Ap
45	26	70.3	200	15 US-10-156-761-8741	

ALIGNMENTS

RESULT 1
US-10-032-950-3
Sequence 3, Application US/10032950
Publication No. US20020077453A1
GENERAL INFORMATION:
APPLICANT: Mair, Tom
APPLICANT: Mayville, Patricia
APPLICANT: No. US20020077453A1ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: Ji, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/10/032,950
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-3

Query Match 94.6%; Score 35; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 GVAAXSLF 9
1 GVAAXSLF 9


```
RESULT 2
US-10-156-761-9222
; Sequence 9222, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMTURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9222
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9222

Query Match      86.5%; Score 32; DB 15; Length 391;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVAAXSSLF 9
DB 274 GVAASALF 282

RESULT 3
US-10-080-170-155
; Sequence 155, Application US/10080170
; Publication No. US20030123601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSIS
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 155
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-080-170-155

Query Match      83.8%; Score 31; DB 16; Length 163;
Best Local Similarity 77.8%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSSLF 9
DB 132 GVAASALF 140

RESULT 4
US-10-080-170-538
; Sequence 538, Application US/10080170
; Publication No. US20030129601A1
```

```
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSIS
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 538
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-538

Query Match      83.8%; Score 31; DB 16; Length 167;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSSLF 9
DB 136 GVAASALF 144

RESULT 5
US-10-267-756-2
; Sequence 2, Application US/10267756
; Publication No. US20030235811A1
; GENERAL INFORMATION:
; APPLICANT: Redindo, Matthew
; APPLICANT: Sompop, Bencharit
; APPLICANT: Potter, Philip
; APPLICANT: Norton, Christopher
; TITLE OF INVENTION: CRYSTALLIZED MAMMALIAN CARBOXYLESTERASE POLYPEPTIDE AND SCREEN
; FILE REFERENCE: Attorney Docket No. US20030235811A1 421-63-2
; CURRENT APPLICATION NUMBER: US/10/267,756
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/374,513
; PRIOR FILING DATE: 2002-04-22
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-267-756-2

Query Match      81.1%; Score 30; DB 12; Length 565;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSSLF 9
DB 248 GVALLSLF 256

RESULT 6
US-10-032-950-1
; Sequence 1, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Muir, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
```

;; CURRENT APPLICATION NUMBER: US/10/032,950
;; CURRENT FILING DATE: 2001-12-27
;; PRIOR APPLICATION NUMBER: 60/090,402
;; PRIOR FILING DATE: 1998-06-24
;; NUMBER OF SEQ ID NOS: 8
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 1
;; LENGTH: 9
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;; NAME/KEY: VARIANT
;; LOCATION: (5)
;; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-1

Query Match 78.4%; Score 29; DB 13; Length 9;
Best Local Similarity 88.9%; Pred. No. 7e+05; 1; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 1;

QY 1 GVAAXSLF 9
Db 1 GVAAXSLF 9

RESULT 7
US-09-966-546-10
;; Sequence 10, Application US/09966546
;; Patent No. US20020168716A1
;; GENERAL INFORMATION:
;; APPLICANT: Vernet, Corine
;; APPLICANT: Shimkets, Richard A.
;; TITLE OF INVENTION: No. US20020168716A1 Human Proteins and Polynucleotides Encoding
;; FILE REFERENCE: Cura-46 (15966-546)
;; CURRENT APPLICATION NUMBER: US/09/966,546
;; CURRENT FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: 09/544,511
;; PRIOR FILING DATE: 2000-04-06
;; NUMBER OF SEQ ID NOS: 57
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 10
;; LENGTH: 160
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-966-546-10

Query Match 78.4%; Score 29; DB 10; Length 160;
Best Local Similarity 77.8%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
Db 5 GVAAGSLF 13

RESULT 8
US-09-966-545-10
;; Sequence 10, Application US/09966545
;; Patent No. US20020172999A1
;; GENERAL INFORMATION:
;; APPLICANT: Vernet, Corine
;; APPLICANT: Shimkets, Richard A.
;; TITLE OF INVENTION: No. US20020172999A1 Human Proteins and Polynucleotides Encoding
;; FILE REFERENCE: Cura-46 (15966-546)
;; CURRENT APPLICATION NUMBER: US/09/966,545
;; CURRENT FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: 09/544,511

;; PRIOR FILING DATE: 2000-04-06
;; NUMBER OF SEQ ID NOS: 57
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 10
;; LENGTH: 160
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-966-545-10

Query Match 78.4%; Score 29; DB 10; Length 160;
Best Local Similarity 77.8%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
Db 5 GVAAGSLF 13

RESULT 9
US-09-965-212-10
;; Sequence 10, Application US/09965212
;; Publication No. US20030003462A1
;; GENERAL INFORMATION:
;; APPLICANT: Vernet, Corine
;; APPLICANT: Shimkets, Richard A.
;; TITLE OF INVENTION: No. US20030003462A1 Human Proteins and Polynucleotides Encod
;; FILE REFERENCE: Cura-46 (15966-546)
;; CURRENT APPLICATION NUMBER: US/09/965,212
;; CURRENT FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: US/09/544,511
;; PRIOR FILING DATE: 2000-04-06
;; PRIOR APPLICATION NUMBER: USSN 60/128,514
;; PRIOR FILING DATE: 1999-04-09
;; NUMBER OF SEQ ID NOS: 57
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 10
;; LENGTH: 160
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-965-212-10

Query Match 78.4%; Score 29; DB 11; Length 160;
Best Local Similarity 77.8%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVAAXSLF 9
Db 5 GVAAGSLF 13

RESULT 10
US-10-189-940-10
;; Sequence 10, Application US/10189940
;; Publication No. US20030129613A1
;; GENERAL INFORMATION:
;; APPLICANT: Vernet, Corine
;; APPLICANT: Shimkets, Richard
;; APPLICANT: Anderson, David
;; APPLICANT: Padigaru, Muraidhara
;; APPLICANT: Boldog, Ferenc
;; APPLICANT: Li, Li
;; APPLICANT: Shenoy, Suresh
;; APPLICANT: Casman, Stacie
;; APPLICANT: Rastelli, Luca
;; TITLE OF INVENTION: No. US20030129613A1 Human Proteins and Polynucleotides Encod
;; FILE REFERENCE: 15966-546 CIP
;; CURRENT APPLICATION NUMBER: US/10/189,940
;; CURRENT FILING DATE: 2002-07-03
;; PRIOR APPLICATION NUMBER: 60/303,241
;; PRIOR FILING DATE: 2001-07-05

```

; PRIOR APPLICATION NUMBER: 60/369,065
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/378,730
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 09/965,212
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/966,545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/966,546
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/186,592
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: Curaseqdist version 0.1
; SEQ ID NO 10
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-189-940-10

```

```

Query Match      78.4%; Score 29; DB 16; Length 160;
Best Local Similarity 77.8%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 GVAAXSLF 9
Db      5 GVAAGSLF 13

```

```

RESULT 11
US-09-966-546-12
; Sequence 12, Application US/09966546
; Patent No. US20020168716A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20020168716A1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/966,546
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-966-546-12

```

```

Query Match      78.4%; Score 29; DB 10; Length 161;
Best Local Similarity 77.8%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 GVAAXSLF 9
Db      5 GVAAGSLF 13

```

```

RESULT 12
US-09-966-545-12
; Sequence 12, Application US/09966545
; Patent No. US20020172999A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard A.

```

```

; TITLE OF INVENTION: No. US20020172999A1 Human Proteins and Polynucleotides Encod
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/966,545
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-966-545-12

```

```

Query Match      78.4%; Score 29; DB 10; Length 161;
Best Local Similarity 77.8%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 GVAAXSLF 9
Db      5 GVAAGSLF 13

```

```

RESULT 13
US-09-965-212-12
; Sequence 12, Application US/09965212
; Publication No. US2003003462A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US2003003462A1 Human Proteins and Polynucleotides Encod
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/965,212
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/544,511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: USN 60/128,514
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-965-212-12

```

```

Query Match      78.4%; Score 29; DB 11; Length 161;
Best Local Similarity 77.8%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 GVAAXSLF 9
Db      5 GVAAGSLF 13

```

```

RESULT 14
US-10-189-940-12
; Sequence 12, Application US/10189940
; Publication No. US20030129613A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shimkets, Richard
; APPLICANT: Anderson, David
; APPLICANT: Padigaru, Muraidhara
; APPLICANT: Bollog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Shency, Suresh
; APPLICANT: Casman, Stacie
; APPLICANT: Rastelli, Luca

```

Job time : 25.5441 secs

```

; TITLE OF INVENTION: No. US20030129613A1el Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: 15966-546 CIP
; CURRENT APPLICATION NUMBER: US/10/189,940
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/303,241
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/369,065
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/378,730
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 09/965,212
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/966,545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/966,546
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/186,592
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: Cuiaseqlist version 0.1
; SEQ ID NO 12
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-189-940-12

```

```

Query Match      78.4%; Score 29; DB 16; Length 161;
Best Local Similarity 77.8%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY      1 GVAAXSSLF 9
Db      5 GVAAGSLRF 13

```

```

RESULT 15
US-10-369-493-1181
; Sequence 1181, Application US/10369493
; Publication No. US2003033675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xiaofeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1181
; LENGTH: 1668
; TYPE: PRT
; ORGANISM: Methanobacterium thermoautotrophicum
; US-10-369-493-1181

```

```

Query Match      78.4%; Score 29; DB 12; Length 1668;
Best Local Similarity 66.7%; Pred. No. 8.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

OY      1 GVAAXSSLF 9
Db      1273 GVAISRRLF 1281

```

Search completed: January 29, 2004, 15:01:45


```

XX
PI Muir TW, Mayville P, Novick RP, Ji G, Beavis R,
XX WPI; 2000-147202/13.
XX
XX New cyclic peptides for treating infections with Staphylococcus aureus
PT
XX
PS Claim 9; Page 26; 37pp; English.
XX
XX This sequence represents a cyclic peptide derived from the Staphylococcus
CC aureus AgRD peptide. The invention relates to AgRD derived peptides,
CC a composition containing a peptide and a carrier, and a method for the
CC production of the cyclic peptides. The peptide inhibits the agr response,
CC which is normally associated with the release of virulence factors of
CC Staphylococcus aureus. An AgRD peptide is produced by S. aureus that
CC activates the agr response in strains of a single group, but interferes
CC with this response in strains of different groups. The peptides and
CC composition containing them can be used to treat infections by S. aureus.
XX
SQ Sequence 9 AA;
XX
XX Query Match 94.6%; Score 35; DB 21; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVNAXSLF 9
DB 1 AVNAXSLF 9
XX
XX RESULT 2
XX ABP53543
XX ID ABP53543 standard; peptide; 9 AA.
XX AC ABP53543;
XX DT 13-DEC-2002 (first entry)
XX DE Cyclic peptide SEQ ID NO:4.
XX KM Cyclic peptide; Staphylococcus aureus; infection; antibacterial;
XX KW agr response inhibitor.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 5 /note= "any amino acid"
XX PN US2002077453-A1.
XX PD 20-JUN-2002.
XX PF 27-DEC-2001; 2001US-0032950.
XX PR 24-JUN-1998; 98US-090402P.
XX RR 24-JUN-1999; 99US-0339511.
XX
XX (MUIR/) MUIR T W.
XX (MAYV/) MAYVILLE P.
XX (NOVI/) NOVICK R P.
XX (BEAV/) BEAVIS R.
XX (JIGG/) JI G.
XX
XX Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX WPI; 2002-681366/73.
XX
XX New cyclic peptides, useful for treating Staphylococcus aureus
PT infections
XX
XX Claim 9; Page 10; 18pp; English.

```

```

XX
XX ABP53540 to ABP53547 represent cyclic peptides (I) from the present
CC invention. The present invention also describes a method for treating
CC Staphylococcus aureus infection comprising the administration of a
CC composition comprising (I). (I) has antibacterial activity, and can be
CC used as an agr gene response inhibitor. The peptides are useful for
CC treating S. aureus infections.
XX
SQ Sequence 9 AA;
XX
XX Query Match 94.6%; Score 35; DB 23; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVNAXSLF 9
DB 1 AVNAXSLF 9
XX
XX RESULT 3
XX AAM50902
XX ID AAM50902 standard; Peptide; 9 AA.
XX AC AAM50902;
XX DT 08-MAY-2002 (first entry)
XX DE AgRD-autoinducing cyclic peptide, inhibitor of agr response.
XX KW Staphylococcus aureus; AgRD; agr response; inhibitor; antibiotic;
XX antibacterial; infection; therapy; cyclic.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 5 /note= "any amino acid"
XX FT Misc-difference 5 /note= "note linked to residue 9 to form cyclic
XX peptide"
XX FT Misc-difference 9 /note= "note linked to residue 5 to form cyclic
XX peptide"
XX PN US6337385-B1.
XX PD 08-JAN-2002.
XX PF 24-JUN-1999; 99US-0339511.
XX PR 24-JUN-1998; 98US-090402P.
XX RR (UYRQ ) UNIV ROCKEFELLER.
XX (UYNY ) UNIV NEW YORK STATE.
XX
XX Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX WPI; 2002-170774/22.
XX
XX Novel synthetic, cyclic AgRD-autoinducing peptide for bacterial
PT interference and for treating Staphylococcus aureus infection in a
XX subject
XX
XX Disclosure; Column 17-18; 18pp; English.
XX
XX The present sequence is that of a novel synthetic cyclic peptide
CC of the invention that is capable of inhibiting the agr response of
CC Staphylococcus aureus. It is an AgRD-autoinducing peptide, where
CC AgRD is a secreted agr-encoded peptide and where the agr locus
CC controls the synthesis of virulence factor and other extracellular
CC proteins responsible for pathogenicity in S. aureus. Preferred
CC peptides may have the sequence NH2-X(n)-Z-X(y)-COOH, with a
CC cyclic bond between the Z residue and COOH other than a thioester

```

CC bond, where X is an amino acid, an amino acid analogue, a
 CC peptidomimetic or non-amide isostere, Z is a synthetic or a
 CC biosynthetic amino acid, n is 0-10 and y is 1-10. The cyclic bond
 CC is especially a lactam or lactone bond. The thiololactone
 CC structure within native AgrD peptides is required for activation
 CC of the agr response. Elimination of the thiol ester component of
 CC the cyclic ring structure can destroy agr response activating. A
 CC activity while preserving and enhancing inhibitory activity. A
 CC claimed method of preparing a cyclic peptide involves: assembling
 CC a linear peptide chain on to a solid phase resin support;
 CC deprotecting the resulting protected assembled peptide; treating the
 CC deprotected peptide with neutral buffer for a time sufficient to
 CC form the cyclic peptide and cleave the peptide from the support;
 CC and recovering the cyclic peptide. The peptide is useful for
 CC bacterial interference, especially for the treatment of *S. aureus*
 CC infection.

XX Sequence 9 AA;

Query Match 94.6%; Score 35; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNAXSSSLF 9
 |||||
 1 VNAXSSSLF 9

Db 1 VNAXSSSLF 9

RESULT 4
 AAY67859
 ID AAY67859 standard; peptide; 9 AA.
 XX
 AC AAY67859;
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE Staphylococcus aureus AgrDII derived peptide sequence.
 XX
 KM Staphylococcus aureus infection; treatment; AgrD; agr response;
 KM virulence factor.
 XX
 OS Staphylococcus aureus.
 XX
 PN WO9967286-A2.
 XX
 PD 29-DEC-1999.
 XX
 PF 24-JUN-1999; 99WO-US14562.
 XX
 PR 24-JUN-1998; 98US-0103438.
 XX
 PA (UYRQ) UNIV ROCKEFELLER.
 PA (UYNY) UNIV NEW YORK STATE.
 XX
 PI Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
 XX
 DR WPI; 2000-147202/13.
 XX
 PT New cyclic peptides for treating infections with Staphylococcus aureus
 XX
 PS Examples; Page 22; 37pp; English.

XX This sequence represents the Staphylococcus aureus AgrDII derived
 CC peptide. The invention relates to AgrD derived cyclic peptides, a
 CC composition containing a peptide and a carrier, and a method for the
 CC production of the cyclic peptides. The peptide inhibits the agr response,
 CC which is normally associated with the release of virulence factors of
 CC Staphylococcus aureus. An AgrD peptide is produced by *S. aureus* that
 CC activates the agr response in strains of a single group, but interferes
 CC with this response in strains of different groups. The peptides and
 CC composition containing them can be used to treat infections by *S. aureus*.

SQ Sequence 9 AA;

Query Match 86.5%; Score 32; DB 21; Length 9;
 Best Local Similarity 87.5%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSSSLF 9
 |||||
 2 VNAXSSSLF 9

Db 2 VNAXSSSLF 9

RESULT 5
 AAY67860
 ID AAY67860 standard; peptide; 9 AA.
 XX
 AC AAY67860;
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE Staphylococcus aureus AgrDII derived peptide sequence.
 XX
 KM Staphylococcus aureus infection; AgrD; agr response; treatment;
 KM virulence factor.
 XX
 OS Staphylococcus aureus.
 XX
 PN WO9967286-A2.
 XX
 PD 29-DEC-1999.
 XX
 PF 24-JUN-1999; 99WO-US14562.
 XX
 PR 24-JUN-1998; 98US-0103438.
 XX
 PA (UYRQ) UNIV ROCKEFELLER.
 PA (UYNY) UNIV NEW YORK STATE.
 XX
 PI Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
 XX
 DR WPI; 2000-147202/13.
 XX
 PT New cyclic peptides for treating infections with Staphylococcus aureus
 XX
 PS Examples; Page 22; 37pp; English.

XX This sequence represents the Staphylococcus aureus AgrDII derived
 CC peptide. The invention relates to AgrD derived cyclic peptides, a
 CC composition containing a peptide and a carrier, and a method for the
 CC production of the cyclic peptides. The peptide inhibits the agr response,
 CC which is normally associated with the release of virulence factors of
 CC Staphylococcus aureus. An AgrD peptide is produced by *S. aureus* that
 CC activates the agr response in strains of a single group, but interferes
 CC with this response in strains of different groups. The peptides and
 CC composition containing them can be used to treat infections by *S. aureus*.

SQ Sequence 9 AA;

Query Match 86.5%; Score 32; DB 21; Length 9;
 Best Local Similarity 87.5%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSSSLF 9
 |||||
 2 VNAXSSSLF 9

Db 2 VNAXSSSLF 9

RESULT 6
 AAM50907
 ID AAM50907 standard; peptide; 9 AA.
 XX
 AC AAM50907;
 XX

DT 08-MAY-2002 (first entry)
 XX Protected peptide used in cyclic peptide production.
 DE Staphylococcus aureus; AgrD; agr response; inhibitor; antibiotic;
 KW antibacterial; infection; therapy; cyclic.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "Z-Gly"
 FT Modified-site 5
 FT /note= "Ser(Cbu)"
 FT Modified-site 6
 FT /note= "Ser(Bz1)"
 FT Modified-site 7
 FT /note= "Ser(Bz1)"
 FT Misc-difference 5
 FT /note= "note linked to residue 9 to form cyclic peptide"
 FT Misc-difference 9
 FT /note= "note linked to residue 5 to form cyclic peptide"
 FT
 FT US6337385-B1.
 XX
 PD 08-JAN-2002.
 XX
 PF 24-JUN-1999; 99US-0339511.
 XX
 PR 24-JUN-1998; 98US-090402P.
 XX
 PA (UYRO) UNIV ROCKEFELLER.
 PA (UYNY) UNIV NEW YORK STATE.
 XX
 PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
 XX
 DR WPI; 2002-170774/22.
 XX
 PT Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
 PT interference and for treating Staphylococcus aureus infection in a
 PT subject -
 XX
 PS Disclosure; Column 14; 18pp; English.
 XX
 CC The present sequence is that of a protected peptide used in an
 CC example of the preparation of novel synthetic cyclic peptides of
 CC the invention (see AM5089-906). The peptide corresponds to the
 CC Staphylococcus aureus AgrDI sequence with a Cys to Ser mutation
 CC (lactone). It was synthesized on a Wang-resin using an Fmoc
 CC N-alpha protection strategy. Following chain assembly, the peptide
 CC was cleaved from the support and the Ser-5 residue deprotected by
 CC treatment with a trifluoroacetic acid:anisole:water mixture
 CC (90:5:5) for 4 hr. The partially protected peptide-alpha
 CC carboxylates were then dissolved in DMF and treated with PyBOP
 CC and a catalytic amount of dimethylamino pyridine. Cyclization was
 CC complete after 2 hr. The remaining protecting groups were then
 CC removed by treatment with HF and the peptide purified by HPLC.
 CC The cyclic peptide is capable of inhibiting the agr response of
 CC Staphylococcus aureus. The thiolactone structure within native
 CC AgrD peptides is required for activation of this response.
 CC Replacement of the thiol ester component of the cyclic ring
 CC structure with a lactone (as in the present case) or a lactam can
 CC destroy agr response activating activity while preserving and
 CC enhancing inhibitory activity. The cyclic peptides are useful for
 CC bacterial interference, especially for the treatment of S. aureus
 CC infection.
 CC
 SQ Sequence 9 AA;
 Query Match 86.5%; Score 32; DB 23; Length 9;
 Best Local Similarity 87.5%; Pred. No. 9.3e+05;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 VMAXSPLF 9
 DB 2 VMAXSPLF 9
 RESULT 7
 AAM51002
 ID AAM51002 standard; Peptide; 9 AA.
 XX
 XX AC AAM51002;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE AgrD2 linear thioester peptide.
 XX
 KW AgrD2; agr response; inhibitor; antibiotic; antibacterial;
 KW infection; therapy.
 XX
 OS Staphylococcus aureus.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 9
 FT /note= "C-terminal thioester"
 FT
 FT US6337385-B1.
 XX
 PN 08-JAN-2002.
 XX
 PD 24-JUN-1999; 99US-0339511.
 XX
 PF 24-JUN-1998; 98US-090402P.
 XX
 PR 24-JUN-1998; 98US-090402P.
 XX
 PA (UYRO) UNIV ROCKEFELLER.
 PA (UYNY) UNIV NEW YORK STATE.
 XX
 PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
 XX
 DR WPI; 2002-170774/22.
 XX
 PT Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
 PT interference and for treating Staphylococcus aureus infection in a
 PT subject -
 XX
 PS Example 1; Column 9; 18pp; English.
 XX
 CC The present sequence is that of a novel synthetic AgrD2 linear
 CC thioester peptide. The peptide is derived from the cyclic AgrD2
 CC peptide of Staphylococcus aureus group II. AgrD2 is a secreted
 CC agr-encoded peptide, where the agr locus controls the synthesis of
 CC virulence factor and other extracellular proteins responsible for
 CC pathogenicity in S. aureus. The biological activity of the
 CC synthetic peptide was assayed using cultured S. aureus strains
 CC containing a beta-lactamase reporter gene fused to the agr3
 CC promoter. This allowed activation or inhibition of the agr
 CC response to be monitored spectrophotometrically. Unlike an AgrD2
 CC thiolactone cyclic peptide (see AAM51001), the present peptide
 CC was unable to either activate or inhibit the agr response, even
 CC when added to cultured cells at low concentrations. The invention
 CC provides claimed cyclic peptides (see AAM5089-906 and AAM5099) for
 CC and methods for preparing them. The cyclic peptides are useful for
 CC bacterial interference, especially for the treatment of S. aureus
 CC infection.
 CC
 SQ Sequence 9 AA;
 Query Match 86.5%; Score 32; DB 23; Length 9;
 Best Local Similarity 87.5%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 VMAXSPLF 9

Db 2 VNASSSLF 9

RESULT 8
AAMS1003 standard; Peptide; 9 AA.

AC AAMS1003;

DT 08-MAY-2002 (first entry)

XX AgrD2 linear free acid peptide.

KW AgrD2; agr response; inhibitor; antibiotic; antibacterial;

OS Synthetic.

XX Staphylococcus aureus.

PN US6337385-B1.

PD 08-JAN-2002.

PF 24-JUN-1999; 99US-0339511.

PR 24-JUN-1998; 98US-090402P.

PA (UYRQ) UNIV ROCKEFELLER.

PA (UYNV) UNIV NEW YORK STATE.

PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;

DR WPI; 2002-170774/22.

PT Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial

PT interference and for treating Staphylococcus aureus infection in a

PT subject -

PS Example 1; Column 9; 18pp; English.

XX The present sequence is that of a novel synthetic AgrD2 linear

CC free acid peptide. The peptide is derived from the cyclic AgrD2

CC peptide of Staphylococcus aureus group II. AgrD2 is a secreted

CC agr-encoded peptide, where the agr locus controls the synthesis of

CC virulence factor and other extracellular proteins responsible for

CC pathogenicity in S. aureus. The biological activity of the

CC synthetic peptide was assayed using cultured S. aureus strains

CC containing a beta-lactamase reporter gene fused to the agr3

CC promoter. This allowed activation or inhibition of the agr

CC response to be monitored spectrophotometrically. Unlike an AgrD2

CC thiolactone cyclic peptide (see AAMS1001), the present peptide

CC was unable to either activate or inhibit the agr response, even

CC when added to cultured cells at 10⁶ concentrations. The invention

CC provides claimed cyclic peptides (see AAMS0899-906 and AAMS0999)

CC and methods for preparing them. The cyclic peptides are useful for

CC bacterial interference, especially for the treatment of S. aureus

CC infection.

XX Sequence 9 AA;

SO Query Match 86.5%; Score 32; DB 23; Length 9;

Best Local Similarity 87.5%; Pred. No. 9.3e+05;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSSLF 9

Db 2 VNASSSLF 9

RESULT 9

AAMS1004

ID AAMS1004 standard; Peptide; 9 AA.

XX AAMS1004;

DT 08-MAY-2002 (first entry)

XX AgrD2 lactone cyclic peptide.

KW AgrD2; agr response; inhibitor; antibiotic; antibacterial;

OS Synthetic.

XX Staphylococcus aureus.

PN US6337385-B1.

PD 08-JAN-2002.

PF 24-JUN-1999; 99US-0339511.

PR 24-JUN-1998; 98US-090402P.

PA (UYRQ) UNIV ROCKEFELLER.

PA (UYNV) UNIV NEW YORK STATE.

PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;

DR WPI; 2002-170774/22.

PT Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial

PT interference and for treating Staphylococcus aureus infection in a

PT subject -

PS Example 1; Column 9; 18pp; English.

XX The present sequence is that of a novel synthetic AgrD2

CC lactone cyclic peptide in which residue 5 of the peptide is

CC linked to residue 9 via a lactone bond. The peptide is derived

CC from an AgrD2 peptide of Staphylococcus aureus group II. AgrD2 is

CC a secreted agr-encoded peptide, where the agr locus controls the

CC synthesis of virulence factor and other extracellular proteins

CC responsible for pathogenicity in S. aureus. The biological

CC activity of the synthetic peptide was assayed using cultured S.

CC aureus strains containing a beta-lactamase reporter gene fused to

CC the agr3 promoter. This allowed activation or inhibition of the

CC agr response to be monitored spectrophotometrically. The lactone

CC AgrD2 peptide inhibited the agr response of group I S. aureus

CC strains without activating the agr response in group I, II or III

CC strains. The invention provides claimed cyclic peptides (see

CC AAMS0899-906 and AAMS0999) and methods for preparing them,

CC especially peptides where the cyclic bond is a lactam or lactone

CC bond. The cyclic peptides are useful for bacterial interference,

CC especially for the treatment of S. aureus infection.

XX Sequence 9 AA;

SO Query Match 86.5%; Score 32; DB 23; Length 9;

Best Local Similarity 87.5%; Pred. No. 9.3e+05;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSSLF 9

Db 2 VNASSSLF 9

RESULT 10

ABB07160	ABB07160 standard; peptide; 9 AA.
ID	ABB07160
XX	
AC	ABB07160;
XX	
DT	13-MAR-2002 (first entry)
XX	
DE	Peptide-mediated quorum sensing inhibitor peptide cyclo-XII.
XX	
KW	Autoinducer-2; AI-2; antibiotic; antibacterial; dermatological;
KM	vulnerable; pheromone; agr system; accessory gene regulator; cyclic.
XX	
OS	Synthetic.
XX	
FN	MO200185664-A2.
XX	
PD	15-NOV-2001.
XX	
PF	10-MAY-2001; 2001MO-US15221.
XX	
PR	10-MAY-2000; 2000US-203000P.
XX	
PR	07-DEC-2000; 2000US-254398P.
XX	
PA	(UNPR-) UNIV PRINCETON.
XX	
PA	(QUOR-) QUORX PHARM INC.
XX	
PI	(UYTE-) UNIV TECHNOLOGIES INT INC.
XX	
PI	Bassler BL, Dammel CS, Schauder S, Shokat K, Stein J, Surette MG;
XX	
DR	WPI; 2002-075235/10.
XX	
PT	Use of autoinducer-2 agonists or antagonists for regulating activity of
XX	autoinducer-2 receptor, regulating bacterial growth and pathogenesis,
PT	also antibiotic compositions -
XX	
PS	Disclosure; Page 33; 134pp; English.
XX	
CC	The invention relates to the use of autoinducer-2 (AI-2) agonists or
CC	antagonists for regulating activity of autoinducer-2 receptor, regulating
CC	bacterial growth and pathogenesis. Synergistic antibiotic compositions
CC	comprising inhibitors of the quorum-sensing pathway of a microorganism
CC	are also provided. Methods using such AI-2 analogues are useful for
CC	treating pathogen-associated disease states. The compounds and antibiotic
CC	compositions can be used to inhibit bacterial cell growth and/or biofilm
CC	formation on a medical device, particularly for promoting growth of skin
CC	graft replacements used in the treatment of burns and ulcers. They may
CC	also be used to aid wound repair, and to inhibit bacterial cell growth
CC	and biofilm formation in or on products or devices used for personal
CC	hygiene. The present sequence represents a inhibitor of peptide-mediated
CC	quorum sensing.
XX	
SO	Sequence 9 AA.
XX	
Query Match	86.5%; Score 32; DB 23; Length 9;
Best Local Similarity	87.5%; Pred. No. 9.3e+05;
Matches 7; Conservative	0; Mismatches 1; Indels 0; Gaps 0
OY	2 VNAXSSLF 9
XX	
DB	2 VNAXSSLF 9
XX	
RESULT 11	
ID	AA67851
AA67851	AA67851 standard; peptide; 9 AA.
XX	
AC	AA67851;
XX	
DT	25-APR-2000 (first entry)
XX	
DE	S. aureus peptide #1 used for bacterial interference.
XX	
KM	Staphylococcus aureus infection; cyclic peptide; AgrT; agr response;

KW	virulence factor; treatment.
XX	
OS	Staphylococcus aureus.
FH	Key
FT	Misc-difference 5
FT	/label= Unknown
PN	WO9967286-A2.
XX	
PD	23-DEC-1999.
XX	
PF	24-JUN-1999; 99WO-US14562.
PR	24-JUN-1999; 98US-0103438.
XX	(UTRC) UNIV ROCKEFELLER.
PA	(UTNY) UNIV NEW YORK STATE.
XX	
P1	Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
XX	
DR	WPI; 2000-147202/13.
PT	New cyclic peptides for treating infections with Staphylococcus aureus
FT	-
XX	
PS	Claim 9; Page 26; 37pp; English.
CC	This sequence represents a cyclic peptide derived from the Staphylococcus
CC	aureus AgRD peptide. The invention relates to AgRD derived peptides,
CC	a composition containing a peptide and a carrier, and a method for the
CC	production of the cyclic peptides. The peptide inhibits the agr response,
CC	which is normally associated with the release of virulence factors of
CC	Staphylococcus aureus. An AgRD peptide is produced by S. aureus that
CC	activates the agr response in strains of a single group, but interferes
CC	with this response in strains of different groups. The peptides and
CC	composition containing them can be used to treat infections by S. aureus.
XX	
SQ	Sequence 9 AA;
Query Match	83.8%; Score 21; DB 21; Length 9;
Best Local Similarity	100.0%; Pred. No. 9.3e+05;
Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY	2 VNPKSLSLF 9 2 VNAKXSLF 9
Db	2 VNAKXSLF 9
RESULT 12	
ID	AA67861 standard; peptide; 9 AA.
XX	
AC	AA67861;
XX	
DT	25-APR-2000 ((first entry))
XX	
DE	Staphylococcus aureus AgRDII derived peptide sequence.
XX	
KW	Staphylococcus aureus infection; AgRD; agr response; treatment;
RW	virulence factor.
XX	
OS	Staphylococcus aureus.
XX	
FH	Key
FT	Misc-difference 5
FT	/label= Unknown
XX	
FN	WO9967286-A2.
XX	
PD	23-DEC-1999.
XX	
PF	24-JUN-1999; 99WO-US14562.

XX 24-JUN-1998; 98US-0103438.
 PR (UYRQ) UNIV ROCKEFELLER.
 PA (UYNY) UNIV NEW YORK STATE.
 XX Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
 PI WPI; 2000-147202/13.
 XX New cyclic peptides for treating infections with Staphylococcus aureus
 PT
 PS Examples; Page 22; 37pp; English.
 XX This sequence represents the Staphylococcus aureus AgrDII derived
 CC peptide. The invention relates to AgrD derived cyclic peptides; a
 CC composition containing a peptide and a carrier, and a method for the
 CC production of the cyclic peptides. The peptide inhibits the agr response,
 CC which is normally associated with the release of virulence factors of
 CC Staphylococcus aureus. An AgrD peptide is produced by S. aureus that
 CC activates the agr response in strains of a single group, but interferes
 CC with this response in strains of different groups. The peptides and
 CC composition containing them can be used to treat infections by S. aureus.
 XX
 SQ Sequence 9 AA;
 Query Match 83.8%; Score 31; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 VNAXSSLP 9
 Db 2 VNAXSSLP 9
 RESULT 13
 ABP53540
 ID ABP53540 standard; peptide; 9 AA.
 AC ABP53540;
 XX 13-DEC-2002 (first entry)
 DT Cyclic peptide SEQ ID NO:1.
 DE Cyclic peptide; Staphylococcus aureus; infection; antibacterial;
 XX agr response inhibitor.
 KW Synthetic.
 OS Key Location/Qualifiers
 FH Misc-difference 5 /note= "any amino acid"
 FT
 FT US2002077453-A1.
 PN 20-JUN-2002.
 PD 27-DEC-2001; 2001US-0032950.
 PF 24-JUN-1998; 98US-090402P.
 PR 24-JUN-1999; 99US-0339511.
 XX (MUIR/) MUIR T W.
 PA (MAYV/) MAYVILLE P.
 PA (NOVI/) NOVICK R P.
 PA (BEAV/) BEAVIS R.
 PA (JIGG/) JI G.
 XX Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
 PI WPI; 2002-681366/73.

XX New cyclic peptides, useful for treating Staphylococcus aureus
 PT infections
 PT
 PS Claim 9; Page 10; 18pp; English.
 XX ABP53540 to ABP53547 represent cyclic peptides (I) from the present
 CC invention. The present invention also describes a method for treating
 CC Staphylococcus aureus infection comprising the administration of a
 CC composition comprising (I). (I) has antibacterial activity, and can be
 CC used as an agr gene response inhibitor. The peptides are useful for
 CC creating S. aureus infections.
 XX
 SQ Sequence 9 AA;
 Query Match 83.8%; Score 31; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 VNAXSSLP 9
 Db 2 VNAXSSLP 9
 RESULT 14
 AAM50899
 ID AAM50899 standard; peptide; 9 AA.
 AC AAM50899;
 XX 08-MAY-2002 (first entry)
 DT AgrD-autoinducing cyclic peptide, inhibitor of agr response.
 DE Staphylococcus aureus; AgrD; agr response; inhibitor; antibiotic;
 XX antibacterial; infection; therapy; cyclic.
 KW Synthetic.
 OS Key Location/Qualifiers
 FH Misc-difference 5 /note= "any amino acid"
 FT Misc-difference 5 /note= "note linked to residue 9 to form cyclic
 FT peptide"
 FT Misc-difference 9 /note= "note linked to residue 5 to form cyclic
 FT peptide"
 FT US6337365-B1.
 PN 08-JAN-2002.
 PD 24-JUN-1999; 99US-0339511.
 PF 24-JUN-1998; 98US-090402P.
 PR (UYRQ) UNIV ROCKEFELLER.
 PA (UYNY) UNIV NEW YORK STATE.
 XX Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
 PI WPI; 2002-170774/22.
 XX Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
 PT interference and for treating Staphylococcus aureus infection in a
 PT subject
 PS Claim 7; Column 19; 18pp; English.
 XX The present sequence is that of a novel synthetic cyclic peptide
 CC of the invention that is capable of inhibiting the agr response of
 CC Staphylococcus aureus. It is an AgrD-autoinducing peptide, where

CC Agd is a secreted agr-encoded peptide and where the agr locus
CC controls the synthesis of virulence factor and other extracellular
CC proteins responsible for pathogenicity in *S. aureus*. Preferred
CC peptides may have the sequence NH₂-X(n)-Z-X'(y)-COOH, with a
CC cyclic bond between the Z residue and COOH other than a thioester
CC bond, where X is an amino acid, an amino acid analogue, a
CC peptidomimetic or non-amide isostere, Z is a synthetic or a
CC biosynthetic amino acid, n is 0-10 and y is 1-10. The cyclic bond
CC is especially a lactam or lactone bond. The thiolactone
CC structure within native Agd peptides is required for activation
CC of the agr response. Elimination of the thiol ester component of
CC the cyclic ring structure can destroy agr response activating
CC activity while preserving and enhancing inhibitory activity. A
CC claimed method of preparing a cyclic peptide involves: assembling
CC a linear peptide chain on to a solid phase resin support;
CC deprotecting the resulting protected assembled peptide; treating the
CC deprotected peptide with neutral buffer for a time sufficient to
CC form the cyclic peptide and cleave the peptide from the support;
CC and recovering the cyclic peptide. The peptide is useful for
CC bacterial interference, especially for the treatment of *S. aureus*
CC infection.

Query Match	83.8%	Score 31	DB 23	Length 9
Best Local Similarity	100.0%	Pred. No.	9.3e+05	
Matches	8	Conservative	0	Indels 0
		Mismatches	0	Gaps 0

QY	2	VNAXSSLF	9
Db	2	VNAXSSLF	9

```

RESULT 15
AAM50908
ID AAM50908 standard; Peptide; 9 AA

```

AC	AAM50908;
XX	
DT	08-MAY-2002 (first entry)

DE Protected peptide used in cyclic peptide production.
XX
XX Staphylococcus aureus; Agrd; agr response; inhibitor; antibiotic.
KW antibacterial; infection; therapy; cyclic.
XX
OS Synthetic.

FT	Key	location/Qualifiers
FT	Modified-site	1 /note= "Z-Gly"
FT	Modified-site	5 /label= Dpr (Boc)
FT	Modified-site	6 /note= "Ser(Bal)"
FT	Modified-site	7 /note= "Ser(Bal)"
FT	Misc-difference	5 /note= "note linked to residue 9 to form cyclid peptide"
FT	Misc-difference	9 /note= "note linked to residue 5 to form cyclid peptide"

PN	US6337385-B1.
XX	
PD	08-JAN-2002.
XX	
PE	24-JUN-1999; 99US-0339511
XX	
PR	24-JUN-1998; 98US-090402P
XX	
PA	(UTRQ) UNIV ROCKEFELLER.

PA (UNYNY) UNIV NEW YORK STATE.
XX
XX
PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G
XX
DR MPI, 2002-170774/22.
XX

PT Novel synthetic, cyclic Arg- α -inducing peptide for bacterial
PT interference and for treating *Staphylococcus aureus* infection in a
PT subject -
XX
PS Disclosures: Column 14; 18pp; English.
v

Disclosure; Column 14; 18pp; English.

The present sequence is that of a novel synthetic cyclic peptide used in an example of the preparation of novel synthetic cyclic peptides of the invention (see AIMS089-306). The peptide corresponds to the Staphylococcus aureus Agt01 sequence with a Cys to dimethylprolionic acid (Dpr) mutation (lactam). It was synthesised on a Wang-resin using an Fmoc N-Alpha protection strategy. Following chain assembly, the peptide was cleaved from the support and the Dpr-5 residue deprotected by treatment with a trifluoroacetic acid:anisole:water mixture (90:5:5) for 4 hr. The partially protected peptide:alpha carboxylates were then dissolved in DMF and treated with PyBOP. Cyclization was complete after 2 hr. The remaining protecting groups were removed by treatment with HF and the peptide purified by HPLC. The cyclic peptide is capable of inhibiting the agr response of Staphylococcus aureus. The thiololactone structure within native AgrD peptides is required for activation of this response. Replacement of the thiol ester component of the cyclic ring structure with a lactam (as in the present case) or a lactone can destroy agr response activating activity while preserving and enhancing inhibitory activity. The cyclic peptides are useful for bacterial interference, especially for the treatment of S. aureus infection.

SQ Sequence 9 AA

Query Match	83.8%	Score 31	DB 23	Length 9
Best Local Similarity	100.0%	Pred. No. 9.3e+05		
Matches	8	Conservative	0	Indels 0
		Mismatches	0	Gaps 0

QY	2 VNAXSSLF	5
Db	2 VNAXSSLF	5

Search completed: January 29, 2004, 14:52:12
Job time : 36.0735 secs

Search completed: January 29, 2004, 14:52:12
Job time : 36.0735 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:46:49 ; Search time 11.1176 Seconds
(without alignments)
77.851 Million cell updates/sec

Title: US-10-032-950-4

Perfect score: 37

Sequence: 1 AVNAXSLF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	34	91.9	633	2 T27215	hypothetical prote
2	32	86.5	2055	2 T31617	hypothetical prote
3	31	83.8	413	2 AH2933	N-carbamoyl-beta-a
4	31	83.8	441	2 G98348	n-carbamoyl-beta-a
5	31	83.8	926	2 S48463	SBG24 protein - ye
6	30	81.1	47	2 C89995	Agd protein [impo
7	30	81.1	265	2 S13098	chlorophyll a/b-bi
8	30	81.1	315	2 A96597	hypothetical prote
9	30	81.1	1016	2 T30942	aminopeptidase (BC
10	30	81.1	1016	2 T30943	aminopeptidase (BC
11	29	78.4	271	2 P90350	hypothetical prote
12	29	78.4	382	2 A28067	lysosomal membrane
13	29	78.4	405	2 A60534	P23/LAMP-1 precurs
14	29	78.4	407	2 A30200	120K lysosomal mem
15	29	78.4	546	2 A32260	cholesterol oxidas
16	29	78.4	599	2 D70104	DNA topoisomerase
17	29	78.4	602	2 H86579	DNA gyrase subunit
18	29	78.4	602	2 C72043	DNA gyrase, chain
19	29	78.4	1049	2 T30825	alpha-mannosidase
20	29	78.4	1054	2 A30239	hydroxymethylgluta
21	28	75.7	273	2 AB3093	dehydrogenase/redu
22	28	75.7	273	2 H98193	probable short-cha
23	28	75.7	292	2 AH2082	hypothetical prote
24	28	75.7	333	1 C69812	ferrichrome ABC tr
25	28	75.7	487	2 S57145	hypothetical prote
26	28	75.7	1200	2 AB1343	DNA-directed DNA p
27	27	73.0	76	2 T26894	hypothetical prote
28	27	73.0	195	2 T46291	hypothetical prote
29	27	73.0	251	2 AE1222	cobalamin adenosyl

30	27	73.0	251	2 AH1575	cobalamin adenosyl
31	27	73.0	260	2 B97225	probable endonuclease
32	27	73.0	301	2 E99233	conserved hypothet
33	27	73.0	368	2 AB2215	hypothetical prote
34	27	73.0	379	2 S76029	hypothetical prote
35	27	73.0	492	2 F70339	glycerol kinase -
36	27	73.0	541	2 D96779	probable 3-ketacy
37	27	73.0	570	2 T33320	hypothetical prote
38	27	73.0	573	2 T27578	hypothetical prote
39	27	73.0	726	2 T15810	hypothetical prote
40	27	73.0	756	2 T38683	hypothetical prote
41	27	73.0	760	2 E83610	conserved hypothet
42	27	73.0	842	2 A87557	non-motile and pha
43	27	73.0	842	2 S27533	histidine protein
44	27	73.0	957	2 A47531	glutamy1 aminopept
45	27	73.0	993	2 A58437	probable homeotic

ALIGNMENTS

RESULT 1
T27215
hypothetical protein Y57G11C.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C/Accession: T27215
R/McMurray, A.
submitted to the EMBL Data Library, September 1997
A/Reference number: Z20330
A/Accession: T27215
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-633 <NII>
A/Cross-references: EMBL:Z29281; PIDD:CA816503.1; GSPDB:GN00022; CESP:Y57G11C.1
A/Experimental source: clone Y57G11C
C/Genetics:
A:Gene: CESP:Y57G11C.1
A/Map position: 4
A/Intons: 70/1; 329/3
C/Superfamily: multidrug resistance protein; ATP-binding cassette homology

Query Match
Best Local Similarity 91.9%; Score 34; DB 2; Length 633;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
DB 327 AINAFSLF 335

RESULT 2
T31617
hypothetical protein Y50E8A.m - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C/Accession: T31617
R/Steward, C.
submitted to the EMBL Data Library, September 1999
A/Reference number: Z21047
A/Accession: T31617
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-2055 <WIL>
A/Cross-references: EMBL:AL117200; NID:e1549770; PIDD:CA855056.1; CESP:Y50E8A.m
C/Genetics:
A:Gene: CESP:Y50E8A.m
A/Intons: 273/3; 447/1; 526/1; 735/1; 1247/1; 1418/1; 1494/1; 1753/3

Query Match
Best Local Similarity 86.5%; Score 32; DB 2; Length 2055;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
 |||||
 Db 1751 ALNAXSLF 1759

RESULT 3

AH2933
 N-cardamyl-beta-alanine amidohydrolyase [imported] - Agrobacterium tumefaciens (strain C
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C:Accession: AH2933
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens CS8.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AH2933
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-413 <KOR>
 A:Cross-references: GB:AE008689; PIDN:AL43886.1; PID:g17741433; GSPDB:GN00187
 A:Experimental source: strain CS8 (Dupont)
 C:Genetics:
 A:Gene: amab
 A:Map position: linear chromosome
 C:Superfamily: N-cardamyl-L-amino acid amidohydrolyase

Query Match 83.8%; Score 31; DB 2; Length 413;
 Best Local Similarity 77.8%; Pred. No. 13;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
 |||||
 Db 246 AVNALGSF 254

RESULT 4

G98348
 n-cardamyl-beta-alanine amidohydrolyase PA0444 [imported] - Agrobacterium tumefaciens (S
 C:Species: Agrobacterium tumefaciens
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
 C:Accession: G98348
 R:Goodner, B.; Hinkle, G.; Galtung, S.; Miller, N.; Blanchard, M.; Ourullo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: G98348
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-441 <KUR>
 A:Cross-references: GB:AE007870; PIDN:AAK0313.1; PID:g15160344; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_L_3478
 A:Map position: linear chromosome
 C:Superfamily: N-cardamyl-L-amino acid amidohydrolyase

Query Match 83.8%; Score 31; DB 2; Length 441;
 Best Local Similarity 77.8%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
 |||||
 Db 274 AVNALGSF 282

RESULT 5

S48463
 SEC24 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YIL109C
 C:Species: Saccharomyces cerevisiae
 C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 19-Apr-2002
 C:Accession: S48463
 R:Bowman, S.; Churcher, C.
 submitted to the EMBL Data Library, September 1994
 A:Reference number: S48455
 A:Accession: S48463
 A:Molecule type: DNA
 A:Residues: 1-926 <BOW>
 A:Cross-references: GB:I247047; EMBL:Z38125; NID:g603997; PID:g763237; GSPDB:GN00009;
 C:Genetics:
 A:Gene: SED:SEC24; SEC24; MIPS:YIL109C
 A:Cross-references: MIPS:YIL109C; SGD:S0001371
 A:Map position: 9L
 C:Function:
 A:Description: involved in endoplasmic reticulum to Golgi transport; required for vesi

Query Match 83.8%; Score 31; DB 2; Length 926;
 Best Local Similarity 75.0%; Pred. No. 32;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSLF 9
 |||||
 Db 787 INATSLF 794

RESULT 6

C8995
 Agid protein [imported] - Staphylococcus aureus (strain N315)
 C:Species: Staphylococcus aureus
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:Accession: C8995
 R:Khodur, M.; Ohba, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; C
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
 C.; Shiba, T.; Hatori, M.; Ogasawara, N.; Hayashi, H.; Hiratake, K.
 Lancet 357, 1225-1240, 2001
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A:Reference number: A89758; MUID:21311952; PMID:11438146
 A:Accession: C8995
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-47 <KUR>
 A:Cross-references: GB:BA000018; PID:g13701831; PIDN:BB43124.1; GSPDB:GN00149
 A:Experimental source: strain N315
 C:Genetics:
 A:Gene: agid

Query Match 81.1%; Score 30; DB 2; Length 47;
 Best Local Similarity 87.5%; Pred. No. 2.1;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSLF 9
 |||||
 Db 25 VNACSLF 32

RESULT 7

S13098
 chlorophyll a/b-binding protein precursor - maize
 C:Species: Zea mays (maize)
 C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
 C:Accession: S13098
 R:Virec, J.F.; Schantz, M.L.; Schantz, R.
 Nucleic Acids Res. 18, 7179, 1990
 A:Title: Nucleotide sequence of a maize cDNA coding for a light-harvesting chlorophyll
 A:Reference number: S13098; MUID:91088340; PMID:2263499
 A:Accession: S13098
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-265 <VIR>
 A:Cross-references: EMBL:X55892; NID:g22354; PIDN:CAA39376.1; PID:g22355
 C:Superfamily: Chlorophyll a/b-binding protein

C;Keywords: chloroplast; thylakoid; transmembrane protein

Query Match 81.1%; Score 30; DB 2; Length 265;
Best Local Similarity 77.8%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
DB 17 AVNVPSSLF 25

RESULT 8

A:Accession: A96597

hypothetical protein TSA14.12 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C;Accession: A96597

R;Theologian, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Hultar, L.

Nature 408, 815-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.M.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A66141; PMID:11130712

A;Accession: A96597

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-315 <STO>

A;Cross-references: GB:AE005173; NID:g4204267; PIDN:AD010648.1; GSPDB:GN00141

C;Genetics:

A;Gene: TSA14.12

A;Map position: 1

Query Match 81.1%; Score 30; DB 2; Length 315;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
DB 302 AINAKSLF 310

RESULT 9
T30942
aminopeptidase (EC 3.4.11.-) - Indian meal moth

C;Species: Plodia interpunctella (Indian meal moth)

C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 01-Feb-2002

C;Accession: T30942

R;Zhu, Y.C.; Oppert, B.; Kramer, K.J.; McLaughly, W.H.; Dowdy, A.K.

submitted to the EMBL Data Library, November 1997

A;Description: Molecular comparison of aminopeptidase cDNAs and gene structure between t

A;Reference number: Z20942

A;Accession: T30942

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1016 <ZHU>

A;Cross-references: EMBL:AF034483; NID:g2645992; PID:g645993; PIDN:AC36148.1

C;Superfamily: membrane alanyl aminopeptidase

C;Keywords: aminopeptidase

RESULT 10

T30943

aminopeptidase (EC 3.4.11.-) - Indian meal moth

C;Species: Plodia interpunctella (Indian meal moth)

C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 01-Feb-2002

C;Accession: T30943

R;Zhu, Y.C.; Oppert, B.; Kramer, K.J.; McLaughly, W.H.; Dowdy, A.K.

submitted to the EMBL Data Library, November 1997

A;Description: Molecular comparison of aminopeptidase cDNAs and gene structure between

A;Reference number: Z20942

A;Accession: T30943

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1016 <ZHU>

A;Cross-references: EMBL:AF034484; NID:g2645994; PID:g2645995; PIDN:AC36147.1

C;Superfamily: membrane alanyl aminopeptidase

C;Keywords: aminopeptidase

RESULT 11

F90350

hypothetical protein SSO1873 [imported] - Sulfolobus solfataricus

C;Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C;Accession: F90350

R;She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyaz, M.J.; Ch; Jongs, I.; Jeffries, A.C.; Kozera, C.O.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder; arett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A;Description: Sulfolobus solfataricus complete genome.

A;Reference number: A99139

A;Accession: F90350

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-271 <KOR>

A;Cross-references: GB:AE006641; NID:g13815127; PIDN:AAK42061.1; GSPDB:GN00155

C;Genetics:

A;Gene: SSO1873

Query Match 78.4%; Score 29; DB 2; Length 271;
Best Local Similarity 66.7%; Pred. No. 25;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
DB 249 SVNSDSSLF 257

RESULT 12
A28067
lysosomal membrane glycoprotein LAMP-1 - mouse

C;Species: Mus musculus (house mouse)

C;Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 20-Aug-1999

C;Accession: A28067

R;Chen, J.W.; Cna, Y.; Yukeel, K.V.; Gracy, R.W.; August, J.T.

J. Biol. Chem. 263, 8754-8758, 1988

A;Title: Isolation and sequencing of a cDNA clone encoding lysosomal membrane glycopro

A;Reference number: A28067; PMID:88243732; PMID:3379044

A;Accession: A28067

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-382 <CHP>

A;Cross-references: GB:U03881; NID:g198706; PIDN:AAA39411.1; PID:g293692

A;Note: the authors translated the codon ATT for residue 1 as Leu and CCG for residue

C;Superfamily: lysosome-associated membrane protein

C;Keywords: glycoprotein; membrane protein

Query Match 78.4%; Score 29; DB 2; Length 382;
 Best Local Similarity 75.0%; Pred. No. 37;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSLF 9
 DB 257 MNASSSLF 264

RESULT 13

P2B/LAMP-1 precursor - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 03-May-1996
 C:Accession: A60534
 R:Hefferman, M.; Yousefi, S.; Dennis, J.W.
 Cancer Res. 49, 6077-6084, 1989
 A:Title: Molecular characterization of P2B/LAMP-1, a major protein target of a metastasizing cell line.
 A:Reference number: A60534; MUID:90002989; PMID:2676155
 A:Accession: A60534
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-405 <HEF>
 C:Superfamily: lysosome-associated membrane protein

Query Match 78.4%; Score 29; DB 2; Length 405;
 Best Local Similarity 75.0%; Pred. No. 39;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSLF 9
 DB 280 MNASSSLF 287

RESULT 14

A30200
 120K lysosomal membrane glycoprotein precursor - rat
 N:Alternate names: sialoglycoprotein
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 18-Oct-1989 #sequence_revision 30-Sep-1991 #text_change 20-Aug-1999
 C:Accession: A30200; S03331
 R:Howe, C.L.; Granger, B.L.; Hull, M.; Green, S.A.; Gabel, C.A.; Helenius, A.; Mellman, Proc. Natl. Acad. Sci. U.S.A. 85, 7577-7581, 1988
 A:Title: Derived protein sequence, oligosaccharides, and membrane insertion of the 120-K glycoproteins.
 A:Reference number: A30200; MUID:89017240; PMID:3174652
 A:Accession: A30200

A:Molecule type: mRNA
 A:Residues: 1-407 <HOM>
 A:Cross-references: EMBL:J03672
 A>Note: the authors translated the codon GGC for residue 15 as Val
 R:Himeno, M.; Noguchi, Y.; Sasaki, H.; Tanaka, Y.; Furuno, K.; Kono, A.; Sasaki, Y.; Kat
 FEBS Lett. 244, 351-356, 1989
 A:Title: Isolation and sequencing of a cDNA clone encoding 107 kDa sialoglycoprotein in
 A:Reference number: S03331; MUID:89155589; PMID:2920855
 A:Accession: S03331

A:Molecule type: mRNA
 A:Residues: 22-407 <HIM>
 A:Cross-references: EMBL:X14765; NID:956577; PIDN:CAA32873.1; PID:956578
 A>Note: part of this sequence, including the amino end of the mature protein, was confir
 C:Superfamily: lysosome-associated membrane protein
 C:Keywords: glycoprotein; membrane protein
 F:22-407/Product: lysosomal membrane glycoprotein, 107K #status experimental <NAT>

Query Match 78.4%; Score 29; DB 2; Length 407;
 Best Local Similarity 75.0%; Pred. No. 39;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSLF 9
 DB 282 MNASSSLF 289

RESULT 15

A32260
 cholesterol oxidase (EC 1.1.3.6) precursor (validated) - Streptomyces sp.
 C:Species: Streptomyces sp.
 C>Date: 20-Oct-1989 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
 C:Accession: A32260; S15810; PC2002
 R:Shizaki, T.; Hirayama, N.; Shinkawa, H.; Nimi, O.; Murooka, Y.
 J. Bacteriol. 171, 596-601, 1989
 A:Title: Nucleotide sequence of the gene for cholesterol oxidase from a Streptomyces
 A:Reference number: A32260; MUID:89123081; PMID:2914858
 A:Accession: A32260

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2546 <ISH>
 A:Cross-references: GB:M31939; GB:J03356; NID:9153210; PIDN:AAA26719.1; PID:9153212
 A>Note: the authors translated the initiation codon GTG for residue 1 as Val
 R:Horii, M.; Shizaki, T.; Paik, S.Y.; Manome, T.; Murooka, Y.
 J. Bacteriol. 172, 3644-3653, 1990
 A:Title: An operon containing the genes for cholesterol oxidase and a cytochrome P-450
 A:Reference number: S15809; MUID:90299781; PMID:2361941
 A:Accession: S15810

A:Molecule type: DNA
 A:Residues: 1-30 <HOR>
 A:Cross-references: EMBL:M31939; GB:J03356; NID:9153210
 R:Purcell, J.P.; Greenplate, J.T.; Jennings, M.G.; Ryerse, J.S.; Pershing, J.C.; Sims
 Biochem. Biophys. Res. Commun. 196, 1406-1413, 1993
 A:Title: Cholesterol oxidase: a potent insecticidal protein active against boll weev
 A:Reference number: PC2002; MUID:94071904; PMID:8250897
 A:Accession: PC2002

A:Molecule type: protein
 A:Residues: 'XXXTF', 48-54, 'XX', 57, 'X', 59-60 <PUR>
 C:Genetics:
 A:Gene: choa

C:Keywords: oxidoreductase
 F:1-42/Domain: signal sequence #status predicted <SIG>
 F:43-546/Product: cholesterol oxidase #status experimental <NAT>

Query Match 78.4%; Score 29; DB 2; Length 546;
 Best Local Similarity 66.7%; Pred. No. 54;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
 DB 446 AVNAXSLF 454

Search completed: January 29, 2004, 14:58:22
 Job time : 12.1176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:45:44 ; Search time 5.82353 Seconds
(without alignments)
72.678 Million cell updates/sec

Title: US-10-032-950-4

Perfect score: 37

Sequence: 1 AVNAXSLF 9

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	83.8	926	1	SC24_YEAST
2	30	81.1	265	1	CB29_WAIZE
3	29	81.1	267	1	PROB_STRTR
4	30	78.4	406	1	LMPI_MOUSE
5	29	78.4	407	1	LMPI_RAT
6	29	78.4	546	1	CHOD_SFRSO
7	28	78.4	589	1	PAPB_BORBT
8	28	78.4	1054	1	HMD1_YEAST
9	28	75.7	497	1	Y92_YEAST
10	28	75.7	1200	1	DP3A_CAMU3
11	27	73.0	379	1	Y528_SYNY3
12	27	73.0	492	1	GLPK_AQUAE
13	27	73.0	756	1	YDLF_SCHPO
14	27	73.0	787	1	NADP_DROME
15	27	73.0	842	1	PUBC_CAUCR
16	27	73.0	957	1	AMPB_HUMAN
17	27	73.0	993	1	TSH_DROME
18	27	73.0	1905	1	TAGB_DICDI
19	27	73.0	2116	1	MY52_DICDI
20	26	70.3	168	1	ATPD_HUMAN
21	26	70.3	174	1	BS5_ANASP
22	26	70.3	184	1	LB5_RENRP
23	26	70.3	187	1	YK45_YEAST
24	26	70.3	286	1	NADC_ECOLI
25	26	70.3	338	1	LAID_ENTCL
26	26	70.3	338	1	LAID_PSEFL
27	26	70.3	338	1	LAID_PSESO
28	26	70.3	364	1	FL6P_ALCEU
29	26	70.3	364	1	FL6P_ALCEU
30	26	70.3	379	1	IDHB_CABEL
31	26	70.3	384	1	CYB_CVACA
32	26	70.3	408	1	CSD_MYCPN
33	26	70.3	444	1	YOJA_BACSU

ALIGNMENTS

RESULT 1
SC24_YEAST
ID SC24_YEAST
AC P404E2, STANDARD: PRT, 926 AA.
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Protein transport protein Sec24 (Abnormal nuclear morphology 1).
GN SEC24 OR ANU1 OR YJL109C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCB1_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8288C / AB972;
RX PubMed=9169870;
RA Churcher C.M., Bowden S., Badcock K., Barker A., Brown D.,
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.E., Hornell T., Hunt S., Jagels K., Jones M., Lye G.,
RA Moute S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrett B.G.,
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX";
RT Nature 387:84-87(1997).
RL [2]
RN CHARACTERIZATION.
RX MEDLINE=20177547; PubMed=10712514;
RA Kuritani T., Hamamoto S., Gimeno R.E., Kaiser C.A., Schekman R.,
RA Yoshida T.;
RT "Sec24 and Isip function interchangeably in transport vesicle
RT formation from the endoplasmic reticulum in Saccharomyces
RT cerevisiae";
RT Mol. Biol. Cell 11:983-998(2000).
RL
CC - FUNCTION: COMPONENT OF THE COP1 COAT, THAT COVERS ER-DERIVED
CC VESICLES INVOLVED IN TRANSPORT FROM THE ENDOPLASMIC RETICULUM TO
CC THE GOLGI APPARATUS. COP1 ACTS IN THE CYTOSOL TO PROMOTE THE
CC TRANSPORT OF SECRETORY, PLASMA MEMBRANE, AND VACUOLAR PROTEINS
CC FROM THE ENDOPLASMIC RETICULUM TO THE GOLGI COMPLEX.
CC - SUBUNIT: COP1 IS COMPOSED OF AT LEAST FIVE PROTEINS: THE SEC23/24
CC COMPLEX, THE SEC13/31 COMPLEX AND SAR1.
CC - SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC - SIMILARITY: BELONGS TO THE SEC23/SEC24 FAMILY. SEC24 SUBFAMILY.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

```

CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL/238125; CAA86271.1; -
CC PIR/548463; S48463.
CC PDB/1M2V; 20-SEP-02.
CC SGD/S0001371; SEC24.
CC GO/GO:0005315; F:Protein binding activity; IDA.
CC DR/GO/GO:0006914; P:autophagy; IMP.
CC DR/InterPro:IPR001974; Gelsolin.
CC DR/InterPro:IPR006900; Sec23_helical.
CC DR/InterPro:IPR006896; Sec23_trunk.
CC DR/InterPro:IPR006895; 2f-Sec23_Sec24.
CC DR/Pfam:PF00682; Gelsolin; 1.
CC DR/Pfam:PF04815; Sec23_helical; 1.
CC DR/Pfam:PF04811; Sec23_trunk; 1.
CC DR/Pfam:PF04810; 2f-Sec23_Sec24; 1.
CC KW/Transport; Protein transport; Golgi stack; Endoplasmic reticulum;
CC MultiGene family; 3D-structure.
CC FT/Domain 231 256 ZINC FINGER-LIKE.
CC FT/Domain 110 114 POLY-GLN.
CC FT/Domain 157 160 POLY-PRO.
CC SQ/SEQUENCE 926 AA; 103635 MW; 35E2BD024CCT5899 CRC64;

Query Match 83.8%; Score 31; DB 1; Length 926;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 VNAXSLF 9
Db 787 INATSLF 794

RESULT 2
CB29_MAIZE STANDARD; PRT; 265 AA.
ID CB29_MAIZE
AC P27497;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chlorophyll A-B binding protein M9, chloroplast precursor (LHCII type
DE I CAB-M9) (LHCP).
GN CAB-M9.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCA clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. W22;
RX MEDLINE=9108340; PubMed=2263499;
RA Viret J.F., Schantz M.L., Schantz R.;
RT "Nucleotide sequence of a maize cDNA coding for a light-harvesting
RT chlorophyll a/b binding protein of photosystem II."
RL Nucleic Acids Res. 18:7179-7179(1990)

CC -1- FUNCTION: THE LIGHT-HARVESTING COMPLEX (LHC) FUNCTIONS AS A LIGHT
CC RECEPTOR. IT CAPTURES & DELIVERS EXCITATION ENERGY TO PHOTOSYSTEMS
CC WITH WHICH IT IS CLOSELY ASSOCIATED. THE N-TERMINUS OF THE PROTEIN
CC EXTENDS INTO THE STROMA WHERE IT IS INVOLVED WITH ADHESION OF
CC GRANAL MEMBRANES AND PHOTOREGULATED BY REVERSIBLE PHOSPHORYLATION
CC OF ITS THREONINE RESIDUES. BOTH ARE BELIEVED TO MEDIANE THE
CC DISTRIBUTION OF EXCITATION ENERGY BETWEEN PHOTOSYSTEMS I AND II.
CC -1- SUBUNIT: THE LHC COMPLEX CONSISTS OF CHLOROPHYLLS (A & B) AND
CC -1- CHLOROPHYLL A-B BINDING PROTEINS.
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).

```

```

CC -----
CC EMBL/X55892; CAA9376.1; -
CC PIR/S13098; S13098.
CC DR/MAZEDB; 61648; -
CC DR/InterPro:IPR001344; Chloro_Abbind.
CC Pfam:PF00504; Chloro_a-b-bind; 1.
CC DR/ProDom; PD000275; Chloro_Abbind; 1.
CC KW/Chlorophyll; Photosynthesis; Photosystem I; Photosystem II;
CC Thylakoid; Membrane; Chloroplast; Transl. peptide; MultiGene family;
CC Transmembrane; Phosphorylation.
CC FT/CHAIN 1 31 CHLOROPHYLL (PROBABLY).
CC FT/CHAIN 32 265 CHLOROPHYLL A-B BINDING PROTEIN M9.
CC FT/TRANSMEM 99 118 POTENTIAL.
CC FT/TRANSMEM 151 171 POTENTIAL.
CC FT/TRANSMEM 219 235 POTENTIAL.
CC SQ/SEQUENCE 265 AA; 28041 MW; 1CC147D3044AC4A1 CRC64;

Query Match 81.1%; Score 30; DB 1; Length 265;
Best Local Similarity 77.8%; Pred. No. 7.6;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AVNAXSLF 9
Db 17 AVNAPSLF 25

RESULT 3
PROB_STRTR STANDARD; PRT; 267 AA.
ID PROB_STRTR
AC P36488;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutamate 5-kinase (EC 2.7.2.11) (Gamma-glutamyl kinase) (GK).
GN PROB.
OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1308;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19258;
RX MEDLINE=97124211; PubMed=8969524;
RA Limauro D., Falcatore A., Basco A.L., Forlani G., de Felice M.;
RT "Proline biosynthesis in Streptococcus thermophilus: characterization
RT of the proB operon and its products."
RL Microbiology 142:3275-3282(1996).
CC -1- FUNCTION: Catalyzes the transfer of a phosphate group to glutamate
CC to form glutamate 5-phosphate which rapidly cyclizes to 5-
CC oxoproline.
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate = ADP + L-glutamate 5-
CC phosphate.
CC -1- PATHWAY: Proline biosynthesis; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: Belongs to the glutamate 5-kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL/X92418; CAA63147.1; -
CC DR/HAMAP; MF_00456; -; 1.
CC DR/InterPro:IPR001048; Aa_kinase.
CC DR/InterPro:IPR001057; Gln_5Kinase.
CC Pfam:PF00696; aakkinase; 1.
CC DR/PRINTS; PR00474; GLUTKINASE.
CC DR/PROSITE; PS00902; GLUTAMATE_5_KINASE; 1.
CC KW/transferase; Kinase; Proline biosynthesis.
CC SQ/SEQUENCE 267 AA; 28966 MW; B7244BC2B7432EB3 CRC64;

```

Query Match 81.1%; Score 30; DB 1; Length 267;
 Best Local Similarity 66.7%; Pred. No. 7.7;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
 DB 121 AINAFESLF 129

RESULT 4
 ID LAMP1_MOUSE STANDARD; PRT; 406 AA.
 AC P14562; 062020; (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 15, Last sequence update)
 DT 01-AUG-1990 (Rel. 15, Last annotation update)
 DE Lysosome-associated membrane glycoprotein 1 precursor (LAMP-1) (LGP-A)
 DE (LGP-120) (CD107A) (P23).
 GN LAMP1 OR LAMP-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90307738; PubMed=2142158;
 RA Grainger B.L., Green S.A., Gabel C.A., Howe C.L., Mellman I.,
 RA Helentius A.;
 RT "Characterization and cloning of lgpl10, a lysosomal membrane
 RT glycoprotein from mouse and rat cells."
 RL J. Biol. Chem. 265:12035-12043(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Heffernan M., Yousefi S., Dennis J.W.;
 RL Submitted (Feb-1990) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 25-406 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=88243732; PubMed=3379044;
 RA Chen J.W., Cha Y., Yaksel K.U., Gracy R.W., August J.T.;
 RT "Isolation and sequencing of a cDNA clone encoding lysosomal membrane
 RT glycoprotein mouse LAMP-1. Sequence similarity to proteins bearing
 RT once-differentiation antigens."
 RL J. Biol. Chem. 263:8754-8758(1988).
 RN [4]
 RP DISULFIDE BONDS.
 RX MEDLINE=90237040; PubMed=2332434;
 RA Arterburn L.M., Earles B.J., August J.T.;
 RT "The disulfide structure of mouse lysosome-associated membrane
 RT protein 1."
 RL J. Biol. Chem. 265:7419-7423(1990).
 CC -1- FUNCTION: PRESENTS CARBOHYDRATE LIGANDS TO SELECTINS. ALSO
 CC IMPLICATED IN TUMOR CELL METASTASIS.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Lysosomal.
 CC THIS PROTEIN SHUTTLES BETWEEN LYSSOMES, ENDOSOMES, AND THE PLASMA
 CC MEMBRANE.
 CC -1- PTM: O- AND N-GLYCOSYLATED. SOME OF THE N-GLYCANS ATTACHED TO
 CC LAMP-1 ARE POLYLACTOSAMINOGLYCANS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE LAMP FAMILY.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M32015; AAA39428.1;
 DR EMBL: M35244; AAA39869.1;
 DR EMBL: J03861; AAA39411.1;
 DR FIR: A28067; A28067.
 DR FIR: A60534; A60534.

DR MGD; MGI:96745; Lamp1.
 DR InterPro; IPR002000; Lamp.
 DR Pfam; PF01299; Lamp. 1.
 DR PRINTS; PR00336; LYSASOCTDMP.
 DR PROSITE; PS00310; LAMP 1; 2.
 DR PROSITE; PS00311; LAMP 2; 1.
 KW Transmembrane; Glycoprotein; Lysosome; Signal.
 FT SIGNAL 1
 FT CHAIN 25 406
 FT FT
 FT DOMAIN 25 370
 FT TRANSMEM 371 394
 FT DOMAIN 395 406
 FT DOMAIN 25 188
 FT DOMAIN 189 218
 FT DOMAIN 219 370
 FT DISULFID 35 74
 FT DISULFID 149 185
 FT DISULFID 222 259
 FT DISULFID 327 364
 FT CARBOHYD 31 31
 FT CARBOHYD 52 32
 FT CARBOHYD 58 58
 FT CARBOHYD 58 58
 FT CARBOHYD 70 70
 FT CARBOHYD 78 78
 FT CARBOHYD 97 97
 FT CARBOHYD 101 101
 FT CARBOHYD 115 115
 FT CARBOHYD 159 159
 FT CARBOHYD 177 177
 FT CARBOHYD 214 214
 FT CARBOHYD 219 219
 FT CARBOHYD 232 232
 FT CARBOHYD 240 240
 FT CARBOHYD 252 252
 FT CARBOHYD 282 282
 FT CARBOHYD 296 296
 FT CARBOHYD 311 311
 FT CONFLICT 1 10
 FT CONFLICT 25 26
 FT CONFLICT 385 385
 SQ SEQUENCE 406 AA; 43865 MM; C1BD73548BB955 CRC64;

Query Match 78.4%; Score 29; DB 1; Length 406;
 Best Local Similarity 75.0%; Pred. No. 21;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSLF 9
 DB 281 MNAXSLF 288

RESULT 5
 ID LAMP1_RAT STANDARD; PRT; 407 AA.
 AC P14562; P97620; (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-JAN-1990 (Rel. 13, Last annotation update)
 DE Lysosome-associated membrane glycoprotein 1 precursor (LAMP-1) (L20
 DE kDa lysosomal membrane glycoprotein) (LGP-120) (CD107A).
 GN LAMP1 OR LAMP-1.
 GN Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86017240; PubMed=3174652;
 RA Howe C.L., Grainger B.L., Hull M., Green S.A., Gabel C.A., Helentius A.,
 RA Mellman I.;
 RT "Derived protein sequence, oligosaccharides, and membrane insertion
 RT of the 120-kDa lysosomal membrane glycoprotein (lgpl120);

RT identification of a highly conserved family of lysosomal membrane
 RT glycoproteins.";
 RT Proc. Natl. Acad. Sci. U.S.A. 85:7577-7581(1988).
 RL [2]
 RN MEDLINE=89153580; PubMed=2920835;
 RP Himeño M., Noguchi Y., Sasaki H., Tanaka Y., Furuno K., Kono A.,
 RA Sakaki Y., Kato K.;
 RT "Isolation and sequencing of a cDNA clone encoding 107 kDa
 RT sialoglycoprotein in rat liver lysosomal membranes.";
 RL FEBS Lett. 244:351-356(1989).
 RN [3]
 RP SEQUENCE OF 283-357 FROM N.A.
 RC STRAIN=Mistar Kyoto; TISSUE=heart smooth muscle;
 RA Adams L.A., Werny I., Schwartz S.M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PRESENTS CARBOHYDRATE LIGANDS TO SELECTINS. ALSO
 CC IMPLICATED IN TUMOR CELL METASTASIS.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein, lysosomal.
 CC THIS PROTEIN SHUTTLES BETWEEN LYOSOMES, ENDOSOMES, AND THE PLASMA
 CC MEMBRANE.
 CC -1- PTM: O- AND N-GLYCOSYLATED, SOME OF THE N-GLYCANS ATTACHED TO
 CC LAMP-1 ARE POLYACTOSAMINOGLYCANS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE LAMP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M34959; AAA4525.1; -;
 DR EMBL; X14765; CA32873.1; -;
 DR EMBL; U75406; AAB19108.1; -;
 DR PIR; A30200; A30200.
 DR InterPro; IPRO02000; LAMP.
 DR Pfam; PF01299; LAMP; 1.
 DR PRINTS; PR00336; LYSASOCTDMP.
 DR PROSITE; PS00310; LAMP_1; 2.
 DR PROSITE; PS00311; LAMP_2; 1.
 DR Transmembrane; Glycoprotein; Lysosome; Signal.
 KM SIGNAL 1 21
 FT CHAIN 22 407
 FT 1. LYSOSOME-ASSOCIATED MEMBRANE GLYCOPROTEIN
 FT 1. LUMENAL (POTENTIAL).
 FT TRANSMEM 372 371
 FT DOMAIN 396 407
 FT 396 407 POTENTIAL. CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 22 189
 FT 22 189 FIRST LUMENAL DOMAIN.
 FT DOMAIN 190 219
 FT 190 219 HINGE.
 FT DISULFID 36 75
 FT 36 75 SECOND LUMENAL DOMAIN.
 FT DISULFID 150 186
 FT 150 186 BY SIMILARITY.
 FT DISULFID 223 260
 FT 223 260 BY SIMILARITY.
 FT DISULFID 328 365
 FT 328 365 BY SIMILARITY.
 FT CARBOHYD 32 32
 FT 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 59 59
 FT 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 71 71
 FT 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 79 79
 FT 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 102 102
 FT 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 116 116
 FT 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 125 125
 FT 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 145 145
 FT 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 160 160
 FT 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178
 FT 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 215 215
 FT 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 220 220
 FT 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 233 233
 FT 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 241 241
 FT 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 271 271
 FT 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 283 283
 FT 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 297 297
 FT 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 293 294 GV -> EF (IN REF. 3).
 FT CONFLICT 329 329 N -> T (IN REF. 3).
 FT CONFLICT 356 357 SD -> VT (IN REF. 3).
 SQ SEQUENCE 407 AA; 43969 MW; 25947490749A7C8 CRC64;
 Query Match 78.4%; Score 29; DB 1; Length 407;
 Best Local Similarity 75.0%; Pred. No. 21;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 VNAXSLF 9
 Db 282 VNAXSLF 289
 RESULT 6
 ID CHOD STRSQ STANDARD; PRT; 546 AA.
 AC P12676;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cholesterol oxidase precursor (EC 1.1.3.6) (CHOD).
 GN CHOA.
 OS Streptomyces sp. (strain SA-COO).
 OC Bacteria; Actinobacteria; Actinobacteriales; Streptomycetales;
 OC Streptomycetaceae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1931;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89123081; PubMed=2914858;
 RA Ishikaki T., Hirayama N., Shinkawa H., Nimi O., Murooka Y.;
 RT "Nucleotide sequence of the gene for cholesterol oxidase from a
 RT Streptomyces sp.";
 RL J. Bacteriol. 171:596-601(1989).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS), AND MUTAGENESIS.
 RX MEDLINE=99211873; PubMed=10194345;
 RA Yue O.K., Kass I.U., Sampson N.S., Vrielink A.;
 RT "Crystal structure determination of cholesterol oxidase from
 RT Streptomyces and structural characterization of key active site
 RT mutants.";
 RL Biochemistry 38:4277-4286(1999).
 CC -1- CATALYTIC ACTIVITY: Cholesterol + O(2) = cholesterol-4-en-3-one +
 CC H(2)O(2).
 CC -1- COFACTOR: FAD.
 CC -1- PATHWAY: CHOLESTEROL METABOLISM.
 CC -1- SUBUNIT: Monomer.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE GMC OXIDOREDUCTASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M31939; AAA26719.1; -;
 DR PIR; A32260; A32260.
 DR PDB; 1B4V; 06-FEB-99.
 DR PDB; 1B8S; 09-FEB-99.
 DR PDB; 1C80; 10-MAR-99.
 DR PDB; 1CC2; 11-MAR-99.
 DR PDB; 1JUH; 28-DEC-01.
 DR PDB; 1MKT; 25-FEB-03.
 DR InterPro; IPRO00172; GMC_oxred.
 DR Pfam; PF05159; GMC_oxred_C; 1.
 DR TIGRAME; TIGR01409; Tat_signal_seq; 1.
 DR PROSITE; PS00623; GMC_OXRED_1; 1.
 DR PROSITE; PS00624; GMC_OXRED_2; FALSE_NEG.


```

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxId=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Rose M., Koester P., Entian K.D.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z49622; CAA89653.1; -
DR PIR; S57145; S57145.
DR TRANSFAC; T03226; -
DR SGD; S0003883; CAF17.
DR InterPro; IPR006222; GCV_T.
DR Pfam; PF01571; GCV_T; 1.
DR Hypothetical protein.
KW SEQUENCE 497 AA; 57055 MW; A9970BBD3BF6019C CRC64;
SQ
Query Match 75.7%; Score 28; DB 1; Length 497;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ANXASLTF 9
DB 245 ANXASLTF 253

RESULT 10
DP3A CAMUE STANDARD; PRT; 1200 AA.
ID DP3A CAMUE
AC Q9PFI9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase III alpha subunit (EC 2.7.7.7).
GN DNAE OR CJO718.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxId=197;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NOTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Kell J.M., Churcher C.,
RA Beaman D., Chillingworth T., Davies R.M., Felwell T., Holtroyd S.,
RA Uggles K., Kariyeh A.V., Moul S., Pallen M.V., Penn C.W.,
RA Quail M.A., Rasthream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrett B.G.;
RL "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RT Nature 403:665-668(2000)
CC -1- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
CC RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
CC THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC THE ALPHA CHAIN IS THE DNA POLYMERASE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA) (N).
CC -1- SUBUNIT: DNA polymerase III contains a core (composed of alpha,
CC epsilon and theta chains) that associates with a tau subunit. This
CC core dimerizes to form the POLIIT complex. PolIIT associates
CC with the gamma complex (composed of gamma, delta, delta', psi and
CC chi chains) and with the beta chain to form the complete DNA
CC polymerase III complex (by similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).

```

```

CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-C FAMILY. DNAB
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL139076; CAB22992.1; -
DR PIR; A81343; A81343.
DR InterPro; IPR004013; BHP_C.
DR InterPro; IPR003141; BHP_N.
DR InterPro; IPR004805; POLC_alpha.
DR Pfam; PF02811; BHP_C; 1.
DR Pfam; PF02231; BHP_N; 1.
DR SMART; SM00481; POLIITC; 1.
DR TRANSFAC; TIGR00594; polC; 1.
KW TRANSFERASE; DNA-directed DNA polymerase; DNA replication;
KW complete proteome.
SQ SEQUENCE 1200 AA; 137340 MW; 17D40CA6623D34FC CRC64;
Query Match 75.7%; Score 28; DB 1; Length 1200;
Best Local Similarity 85.7%; Pred. No. 11e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 NAXSLTF 9
DB 938 NAXSLTF 944

RESULT 11
Y528 SYNY3 STANDARD; PRT; 379 AA.
ID Y528 SYNY3
AC Q5518;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein s110528 (EC 3.4.24.-).
GN S110528.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxId=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RL "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome."
RL DNA Res. 2:153-166(1995).
CC -1- COFACTOR: Binds 1 zinc ion per subunit (by similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M50B.
CC -1- SIMILARITY: Contains 2 CBS domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D64006; BAA10876.1; -
DR PIR; S76029; S76029.
DR MEROPS; M50_upe; -
DR InterPro; IPR000644; CBS_domain.
DR InterPro; IPR001193; Peptidase_M50.
DR InterPro; IPR006025; Zn_Mpeptidase.

```

DR Pfam; PF00571; CBS; 1.
 DR Pfam; PF00163; peptidase_M50; 1.
 DR SMART; SMO0116; CBS; 2.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KM Hypothetical protein; Hydrolase; Metalloprotease; Zinc; Transmembrane;
 KW Repeat; CBS domain; Complete proteome.
 FT TRANSMEM 20 40 POTENTIAL.
 FT TRANSMEM 54 74 POTENTIAL.
 FT TRANSMEM 115 135 POTENTIAL.
 FT TRANSMEM 148 168 POTENTIAL.
 FT TRANSMEM 212 232 POTENTIAL.
 FT DOMAIN 260 310 CBS 1.
 FT DOMAIN 327 375 CBS 2.
 FT METAL 75 75 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT SITE 76 76 BY SIMILARITY.
 FT METAL 79 79 ZINC (CATALYTIC) (BY SIMILARITY).
 SQ SEQUENCE 379 AA; 40465 MW; 1CC325160078ECB CRC64;

Query Match 73.0%; Score 27; DB 1; Length 379;
 Best Local Similarity 75.0%; Pred. No. 59;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VNAXSSLF 9
 Db 331 VNAXSSLF 338

RESULT 12
 GLPK_AQUAE STANDARD; PRT; 492 AA.

ID GLPK_AQUAE
 AC 066746;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glycerol kinase (EC 2.7.1.30) (ATP:glycerol 3-phosphotransferase)
 DE (Glycerokinase) (GK).
 GN GLPK OR AQ 434.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=VFS;
 RC MEDLINE=98196666; PubMed=9537320;
 RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.S., Overbeek R., Sneed M.A., Keller M., Anjay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RT Nature 392:353-358(1998).
 RL Nature 392:353-358(1998).
 CC -1- FUNCTION: KEY ENZYME IN THE REGULATION OF GLYCEROL UPTAKE AND
 METABOLISM.
 CC -1- CATALYTIC ACTIVITY: ATP + glycerol = ADP + glycerol 3-phosphate.
 CC -1- PATHWAY: Glycerol utilization; rate-limiting step.
 CC -1- SIMILARITY: BELONGS TO THE PLOKINASE / GLUCONOKINASE /
 GLYCEROKINASE / XYLLOKINASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; AE000690; AAC06710.1; -
 CC PIR; F70339; F70339.
 CC HAMSP; P08859; IGLC.
 CC HAMAP; MF 00166; -; 1.
 DR InterPro; IPR000577; FGGY_kin.
 DR InterPro; IPR005989; Glycerol_kin.
 DR Pfam; PF00570; FGGY; 1.
 DR Pfam; PF02782; FGGY_C; 1.

DR TIGRfam; TIGR01311; glycerol_kin. 1.
 DR PROSITE; PS00445; FGGY_KINASES_2; 1.
 DR PROSITE; PS00933; FGGY_KINASES_1; 1.
 KM Glycerol metabolism; Transferase; Kinase; ATP-binding;
 KW Complete proteome.
 FT NP BIND 148 160 ATP (PROBABLE).
 SQ SEQUENCE 492 AA; 55297 MW; CE5F0B8FF53B37 CRC64;

Query Match 73.0%; Score 27; DB 1; Length 492;
 Best Local Similarity 75.0%; Pred. No. 78;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VNAXSSLF 9
 Db 211 VNAXSSLF 218

RESULT 13
 YDLF_SCHPO STANDARD; PRT; 756 AA.

ID YDLF_SCHPO
 AC P8128;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein C3A12.15 in chromosome I.
 GN SPAC3A12.15.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21648401; PubMed=11659360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Woods J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moutle S., Mungall K., Murphy L., Niblett D., Odeil C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S.,
 RA Skellton J., Simmonds M., Squares R., Squares K., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voiclaert G., Aert R., Robben J., Grymourez B.,
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Botzlm K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wandut R., Fumelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rocher M., Gallardin C., Tallada V.A., Garçon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Ruvetla J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Snpkovski G.V., Uesery D., Barrett B.G., Nurse P.,
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: TO YEAST YJ1029C AND C.ELEGANS T0565.8.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; Z95395; CAB08743.1; -
 DR PIR; T38683; T38683.

DR GenedB.Spmbe; SPAC3A12.15; --
 DR Pfam; PF04100; VP53_N; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 756 AA; 87505 MW; F45BF6355AF31B48 CRC64;

Query Match 73.0%; Score 27; DB 1; Length 756;
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSSLE 9
 DB 358 AVNALLSYF 366

RESULT 14
 NAD: DROME STANDARD; PRT; 787 AA.
 ID NAD: DROME STANDARD; PRT; 787 AA.
 AC Q9VY40;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative glutamine-dependent NAD(+) synthetase (EC 6.3.5.1) (NAD(+)
 synthase [glutamine-hydrolyzing]).
 GN CG9940.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCB1_TaxID=7227;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Boltschakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
 RA Burks K.C., Busan D.A., Butler H., Chadeu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hristin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenisnson J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasco P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei E., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Wuzni D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skipski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Switzkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhu G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195 (2000).
 CC -1- CATALYTIC ACTIVITY: ATP + deamido-NAD(+) + L-glutamine + H(2)O =
 AMP + diphosphate + NAD(+) + L-glutamate.

CC -1- PATHWAY: NAD biosynthesis.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE NAD
 CC SYNTHETASE FAMILY.
 CC -1- SIMILARITY: Contains 1 CN hydrolase domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE003493; AAF48303.1; --
 DR Flybase; FBgn0030512; CG9940.
 DR InterPro; IPR003694; NAD_synthase.
 DR InterPro; IPR003010; Nlase/CNhydrtse.
 DR Pfam; PF00795; CN hydrolase; 1.
 DR Pfam; PF02540; NAD synthase; 1.
 DR TIGRFAMs; TIGR00552; nade; 1.
 DR PROSITE; PS50263; CN_HYDROLASE; 1.
 KW Hypothetical protein; Ligase; NAD; ATP-binding.
 FT DOMAIN 5 294 CN_HYDROLASE.
 FT NP_BIND 325 787 LIGASE.
 FT ACT_SITE 355 362 ATP (BY SIMILARITY).
 FT ACT_SITE 357 357 BY SIMILARITY.
 SQ SEQUENCE 787 AA; 87614 MW; 0B2B010E327F0C5 CRC64;

Query Match 73.0%; Score 27; DB 1; Length 787;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSSLE 9
 DB 450 AVNALLSYF 458

RESULT 15
 PLEC CAUCR STANDARD; PRT; 842 AA.
 ID PLEC CAUCR STANDARD; PRT; 842 AA.
 AC P37854;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Non-motile and phage-resistance protein (BC 2.7.3.-).
 GN PLEC OR CC2482.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
 OC Caulobacteraceae; Caulobacter.
 NC NCB1_TaxID=155892;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=9313840; PubMed=8421696;
 RA Wang S.P., Sharma P.L., Schoenlein P.V., Ely B.;
 RT "A histidine protein kinase is involved in polar organelle
 RT development in Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:630-634 (1993).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nieman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A.S., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolony J.F., Smit J., Craven M.B., Knouri H., Shetty J., Berry K.,
 RA Ueberlack T., Tran K., Wolf A., Vamathevan J., Emoliava M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).
 CC -1- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM INVOLVED
 CC IN THE REGULATION OF POLAR ORGANELLE DEVELOPMENT. PLEC PROBABLY

```

CC CC FUNCTIONS AS A MEMBRANE-ASSOCIATED PROTEIN KINASE THAT TRANSFERS
CC PHOSPHATE TO A RESPONSE REGULATOR
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- PTM: THE C-TERMINAL IS CAPABLE OF AUTOPHOSPHORYLATION.
CC -1- SIMILARITY: Contains 1 histidine kinase domain.
CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M91449; AAA23052.1; -.
DR EMBL; AE005917; AKX24453.1; -.
DR PIR; A87557; A87557.
DR PIR; S27533; S27533.
DR TIGR; CC2482; -.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003661; His_kinase.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00989; HsKKA; 1.
DR Pfam; PF00512; HsKKA; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HsKKA; 1.
DR SMART; SM00021; PAS; 2.
DR TIGRFAMs; TIGR00229; sensory_box; 2.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS50112; PAS; 1.
KW Sensory transduction; Complete proteome.
KW Phosphorylation; Complete proteome.
FT TRANSMEM 29 50 POTENTIAL.
FT TRANSMEM 283 303 POTENTIAL.
FT TRANSMEM 343 363 POTENTIAL.
FT DOMAIN 318 389 PAS.
FT DOMAIN 607 830 HISTIDINE KINASE.
FT MOD_RES 610 610 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 744 744 A -> V (IN REF. 1).
SQ SEQUENCE 842 AA; 89559 MM; DCFP4F092DE2CBF8 CRG64;

```

Query Match 73.0%; Score 27; DB 1; Length 842;

Best Local Similarity 66.7%; Pred. No. 1,4e+02; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
 ||| | |
 Db 22 AVNAPSQVF 30

Search completed: January 29, 2004, 14:53:10
 Job time : 6.82353 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:46:19 ; Search time 26.8676 Seconds
(without alignments)
86,441 Million cell updates/sec

Title: US-10-032-950-4
Perfect score: 37
Sequence: 1 AVMAXSLF 9

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match	Length	ID	Description
1	34	91.9	633	5 O18227	O18227 caenorhabdi
2	32	86.5	807	5 Q90275	Q90275 caenorhabdi
3	31	83.8	441	16 Q8UB66	Q8UB66 streptomyce
4	31	83.8	549	2 Q9EW96	Q9EW96 streptomyce
5	30	81.1	47	16 O33586	O33586 streptomyce
6	30	81.1	150	3 Q92214	Q92214 coprinus ci
7	30	81.1	240	10 Q9LAK3	Q9LAK3 cryzae sativ
8	30	81.1	315	10 Q92VU6	Q92VU6 arabidopsis
9	30	81.1	335	10 Q94BS2	Q94BS2 arabidopsis
10	30	81.1	335	10 Q8L972	Q8L972 arabidopsis
11	30	81.1	356	10 Q9SRM2	Q9SRM2 arabidopsis
12	30	81.1	358	10 Q9C780	Q9C780 arabidopsis
13	30	81.1	359	10 Q9AT39	Q9AT39 glycine max
14	30	81.1	371	10 Q8GX26	Q8GX26 arabidopsis
15	30	81.1	1016	5 O17485	O17485 plodia inte
16	30	81.1	1016	5 O17484	O17484 plodia inte

17	29	78.4	143	16 O8ERG6	O8ERG6 shewanella
18	29	78.4	189	11 O922T9	O922T9 mus musculus
19	29	78.4	261	12 Q9K1P3	Q9K1P3 myxococcus
20	29	78.4	271	17 Q97X81	Q97X81 sulfolobus
21	29	78.4	406	11 Q8VH34	Q8VH34 mus musculus
22	29	78.4	407	11 Q9DC13	Q9DC13 mus musculus
23	29	78.4	480	10 Q9LEG2	Q9LEG2 lycopersico
24	29	78.4	602	16 Q927J0	Q927J0 chlamydia p
25	29	78.4	684	10 Q9Y2P9	Q9Y2P9 homo sapien
26	29	78.4	1049	3 O13344	O13344 emeticella
27	29	78.4	1566	10 Q9LJL0	Q9LJL0 arabidopsis
28	29	78.4	1604	2 Q9KX99	Q9KX99 rickettsia
29	29	78.4	2555	11 Q912E6	Q912E6 mus musculus
30	29	78.4	2561	11 Q8VIE5	Q8VIE5 mus musculus
31	28	75.7	27	12 Q91UP5	Q91UP5 hepatitis c
32	28	75.7	55	4 Q92661	Q92661 homo sapien
33	28	75.7	111	16 O8PEK1	O8PEK1 xanthomonas
34	28	75.7	133	16 O8PEH6	O8PEH6 xanthomonas
35	28	75.7	153	16 Q9CLW6	Q9CLW6 pasteurella
36	28	75.7	237	16 Q8G5R4	Q8G5R4 bifidobacte
37	28	75.7	273	16 Q8U7T3	Q8U7T3 agrobacteri
38	28	75.7	292	16 Q8YUW7	Q8YUW7 anabaena sp
39	28	75.7	333	16 Q34933	Q34933 bacillus su
40	28	75.7	404	12 Q9W122	Q9W122 rangiferine
41	28	75.7	435	16 Q8DU29	Q8DU29 streptococ
42	28	75.7	543	10 Q9LJL0	Q9LJL0 brassica na
43	28	75.7	554	9 Q8W722	Q8W722 cyanophage
44	28	75.7	638	2 O16655	O16655 paracoccus
45	28	75.7	738	5 Q8XYP6	Q8XYP6 drosophila

ALIGNMENTS

RESULT 1
O18227 PRELIMINARY; PRT: 633 AA.
ID O18227
AC O18227;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Y57G11C.1 protein.
GN Y57G11C.1
GN Caenorhabditis elegans.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9069613; PubMed=9851916;
RA none;
RT "genome sequence of the nematode C.elegans: A platform for
investigating biology".
RL Science 282:2012-2018 (1998)
CC -; SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC EMBL; Z99281; CAB16503.1; -;
DR WormPep; Y57G11C.1; CE14926.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00664; ABC_membrane_1.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transport.
SQ SEQUENCE 633 AA; 71639 MW; 8807414AA9884058 CRC64;
Query Match 91.9%; Score 34; DB 5; Length 633;

Best Local Similarity 77.8%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
DB 327 ALNAXSLF 335

RESULT 2

Q9U275 PRELIMINARY; PRT; 807 AA.

AC Q9U275; PRELIMINARY; PRT; 807 AA.
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Y50E8A.16 protein.
GN Y50E8A.16.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Steward C.A.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2016(1998).
CC 1.- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; AL117200; CAB60586.1; -.
DR WormPep; Y50E8A.16; CE24404.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00664; ABC_membrane.1.
DR Pfam; PF00005; ABC_tran.1.
DR PRODOM; PD000006; ABC_transporter.1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transport.
SQ SEQUENCE 807 AA; 91163 MW; B9ABC59866DF1EF CRC64;

Query Match 86.5%; Score 32; DB 5; Length 807;

Best Local Similarity 77.8%; Pred. No. 63;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
DB 503 ALNAXSLF 511

RESULT 3

Q8UB6 PRELIMINARY; PRT; 441 AA.

AC Q8UB6; PRELIMINARY; PRT; 441 AA.
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE N-cardamoyl-beta-alanine amidohydrolase.
GN AAMB OR ATU3070 OR AGR L 3478.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,

RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,
RA Chapman P., Clendenning U., Deachetage G., Gillet M., Grant C.,
RA Kutyavyn T., Levy R., Li M.-g., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.,
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58."
RL Science 294:2317-2323(2001).
RN [2]

RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quicilo B., Goldman B.S., Gao Y., Askenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Tatchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Clejo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
DR EMBL; AE009237; AAL43886.1; ALT_INIT.
DR EMBL; AE008376; AAK90313.1; -.
DR InterPro; IPR002933; Peptidase_M20.
DR Pfam; PF01546; Peptidase_M20; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 441 AA; 47628 MW; 7AD50CF34F86410F CRC64;

Query Match 83.8%; Score 31; DB 16; Length 441;

Best Local Similarity 77.8%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
DB 274 ALNAXSLF 282

RESULT 4

Q9EM96 PRELIMINARY; PRT; 549 AA.

AC Q9EM96; PRELIMINARY; PRT; 549 AA.
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE P1ME protein.
GN P1ME.
OS Streptomyces natalensis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=68242;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=20547809; PubMed=11094342;
RA Aparicio J.F., Fouces R., Mendes M.V., Oliveira N., Martin J.F.,
RT "A complex multienzyme system encoded by five polyketide synthase
genes is involved in the biosynthesis of the 26-membered polyene
RT macrocyclic pimaricin in Streptomyces natalensis."
RL Chem. Biol. 7:895-905(2000).
DR EMBL; AJ278573; CAC20926.1; -.
DR HSSP; P12676; 1B4V.
DR InterPro; IPR000172; GMC_oxred.
DR InterPro; IPR006311; Tat_signal.
DR TIGRFAIR; TIGR01409; Tat_signal_seq; 1.
DR PROSITE; PS00623; GMC_OXRED; 1.
SQ SEQUENCE 549 AA; 59475 MW; 96759DE780AC5A82 CRC64;

Query Match 83.8%; Score 31; DB 2; Length 549;

Best Local Similarity 77.8%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9

DB 449 AVNAARSLF 457

RESULT 5

033586 PRELIMINARY; PRT; 47 AA.

AC 033586; 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE AGRD (AGRD protein).
 GN AGRD OR SA2037 OR SA1842.1 OR SA8065.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
 OS Staphylococcus aureus (strain N315), and
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=158878, 158879, 1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.aureus; STRAIN=SA502A;
 RX MEDLINE=97342847; PubMed=9197262;
 RA J.G., Beavis R., Novick R.P.;
 RT "Bacterial interference caused by autoinducing peptide variants."
 RL Science 276:2027-2030(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
 RX MEDLINE=21311952; PubMed=11419146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Sekizumi Y., Hirakawa H., Kohara S., Goto S., Yabuzaki J.,
 RA Kameiwa M., Yamashita A., Oshima K., Futuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 RT aureus."
 RL Lancet 357:1225-1240(2001).
 DR EMBL: AF001782; AAB63265.1; -;
 DR EMBL: AP003364; BAB58199.1; -;
 DR EMBL: AP003135; BAB43124.1; -;
 SQ Complete proteome.
 SQ SEQUENCE 47 AA; 5149 MW; 22893764DF54BFA7 CRC64;

Query Match 81.1%; Score 30; DB 16; Length 47;
 Best Local Similarity 87.5%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSLF 9
 DB 25 VNAXSLF 32

RESULT 6

092214 PRELIMINARY; PRT; 150 AA.

AC 092214; 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Galactose binding lectin.
 GN GALT.
 OS Coprinus cinereus (Inky cap fungus).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Agaricales; Psathyrellaceae; Coprinopsis.
 OX NCBI_TaxID=5346;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JRS2;
 RA Cooper D.N.W., Boulianne R.P., Charlton S., Farrell E., Sucher A.,
 RA Lu B.C.;
 RT "Fungal Galactins: Sequence and specificity of two isolectins from

RT Coprinus cinereus";
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U64676; AAB06178.1; -;
 DR InterPro; IPR001079; Galactin.
 DR Pfam; PF00337; Gal-bind lectin; 1.
 DR SMART; SM00276; GLBCT; 1.
 KW Lectin.
 SQ SEQUENCE 150 AA; 16671 MW; 296914448066EEEC CRC64;

Query Match 81.1%; Score 30; DB 3; Length 150;
 Best Local Similarity 77.8%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAARSLF 9
 DB 124 AVNAARSLF 132

RESULT 7

09LKK3 PRELIMINARY; PRT; 240 AA.

AC 09LKK3; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Putative tyrosine phosphatase (Fragment).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhacridae; Oryzae; Oryza.
 OX NCBI_TaxID=4550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=jingxi17, and cv. Japonica; TISSUE=Shoot;
 RA Yu F., Zhang A., Zhang F., Chen S.;
 RT "Rice probable tyrosine phosphatase."
 RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF272976; AAF81796.1; -;
 DR HSP; Q12923; 3PDZ.
 DR Gramene; Q9LKK3; -;
 DR InterPro; IPR001478; PDZ.
 DR InterPro; IPR001440; TPR.
 DR PROSITE; PS50106; PDZ; 1.
 FT NON TER 1
 SQ SEQUENCE 240 AA; 27246 MW; 41A865FCB1751ED CRC64;

Query Match 81.1%; Score 30; DB 10; Length 240;
 Best Local Similarity 66.7%; Pred. No. 56;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAARSLF 9
 DB 227 AVNAARSLF 235

RESULT 8

09ZVU6 PRELIMINARY; PRT; 315 AA.

AC 09ZVU6; 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE TSA14.12 protein.
 GN TSA14.12.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Pederspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
 RA Altati H., Araujo R., Hutzar L., Rowley D., Buehler E., Dunn P.,

RA Gonzalez A., Kremetskaya I., Kim C., Lenz C., Li J., Liu S.,
 RA Lueros S., Schwartz J., Shim P., Toriumi M., Vysotskaya V.S.,
 RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.,
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC005823; AAD10648.1; -
 DR InterPro: IPR001478; TPR.
 DR InterPro: IPR001440; TPR.
 DR PROSITE: PS50106; PDZ; 1.
 SQ SEQUENCE 315 AA; 35272 MW; C8AC19F625B303CD CRC64;

Query Match 81.1%; Score 30; DB 10; Length 315;
 Best Local Similarity 66.7%; Pred. No. 74;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
 DB 302 AINAKSLF 310

RESULT 9
 Q94BS2 PRELIMINARY; PRT; 335 AA.

AC Q94BS2; 01-DEC-2001 (TREMBlrel. 13, Created)
 DT 01-DEC-2001 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical 37.4 kDa protein.
 GN TSA14.12 OR AT1055480.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eumossids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;

RA [1]
 RA SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Dale J.M., Gibson H.A., Goldsmith A.D., Jiang P.X., Lee J.M.,
 RA Quach H.H., Tang C., Toriumi M., Yu G., Bowser L., Carninci P.,
 RA Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamitani A.,
 RA Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shim P., Southwick A., Tracy S.E.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,
 RL "Full length cDNA of gene TSA14.12 (GI:4204267)."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

RA [2]
 RA SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Ergu P., Lee J.M.,
 RA Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G.,
 RA Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shim P.,
 RA Southwick A., Davis R.W., Ecker J.R., Theologis A.,
 RL "Arabidopsis Open Reading Frame (ORF) Clones."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY039926; AAK64030.1; -
 DR EMBL: AY079359; AAL85090.1; -
 DR InterPro: IPR001478; PDZ.
 DR InterPro: IPR001440; TPR.
 DR PROSITE: PS50106; PDZ; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 335 AA; 37410 MW; E83B3BF8EEB39E51 CRC64;

Query Match 81.1%; Score 30; DB 10; Length 335;
 Best Local Similarity 66.7%; Pred. No. 78;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
 DB 322 AINAKSLF 330

RESULT 10
 Q8L972 PRELIMINARY; PRT; 335 AA.

AC Q8L972; 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eumossids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;

RA [1]
 RA SEQUENCE FROM N.A.
 RA Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.,
 RL "Full-length messenger RNA sequences greatly improve genome
 annotation."
 RL Genome Biol. 0:0-0(2002).

RA [2]
 RA SEQUENCE FROM N.A.
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RL "Full-length cDNA from Arabidopsis thaliana."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY088606; AAM6135.1; -
 DR InterPro: IPR001478; PDZ.
 DR InterPro: IPR001440; TPR.
 DR PROSITE: PS50106; PDZ; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 335 AA; 37456 MW; 165924929A8F1BE CRC64;

Query Match 81.1%; Score 30; DB 10; Length 335;
 Best Local Similarity 66.7%; Pred. No. 78;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
 DB 322 AINAKSLF 330

RESULT 11
 Q9SRM2 PRELIMINARY; PRT; 356 AA.

AC Q9SRM2; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE P9F8.5 protein.
 GN P9F8.5.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eumossids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;

RA [1]
 RA SEQUENCE FROM N.A.
 RA STRAIN-cv. Columbia;
 RA Lin Y., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
 RA Rongin C.M., Koo H., Fujii C.Y., Utracki T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fries C.M.,
 RL "Arabidopsis thaliana chromosome III BAC P9F8 genomic sequence."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC009991; AAF01509.1; -
 SQ SEQUENCE 356 AA; 39948 MW; 54DFCD6F2FD075F3 CRC64;

Query Match 81.1%; Score 30; DB 10; Length 356;
 Best Local Similarity 77.8%; Pred. No. 83;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
 DB 42 AYDALSLF 50

```

RESULT 12
OC9C780 PRELIMINARY; PRT; 358 AA.
ID 09C780
AC 09C780
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE Hypothetical 40.2 kDa protein.
GN F11B9.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotsid II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=1130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansgorge W., Unsel'd M.,
RA Faltmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
RA Delbeny M., Boutry M., Grivell L.A., Maché R., Fulgomech P.,
RA De Simone V., Choisme N., Artiguenave F., Robert C., Broillet P.,
RA Winkler P., Catolico L., Weissendbach U., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmbach E., Drzonek H., Ertle H., Jordan N., Bangert S.,
RA Wiedemann R., Kranz H., Voess H., Holland R., Brandt P., Nakamura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppe S., Simonati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,
RA Reichelt J., Scharfe M., Schoen O., Barges M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Otenwaelder B., Duchemin D.,
RA Cooke R., Landie M., Berger-Llauro C., Fumelle B., Masny D.,
RA de Han M., Maarse A.C., Alcaraz J.-P., Cortet A., Casacuberta E.,
RA Monfort A., Argirou A., Flores M., Ligouri R., Vitale D.,
RA Mammucip G., Haase D., Schopf H., Radd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.K., Kaul S., Town C.D., Koo H.L., Talon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Utechtack T., Fujii C.Y., Shen T.P.,
RA Creasy T.H., Haas B., Walts R., Wu D., Peterson U., Van Aken S.,
RA Pai G., Miltscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nieman W.C., Salberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Ideasa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shino S., Tabata S., Wada T.,
RA Metarabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
thaliana.";
RL Nature 408:820-822(2000).
DR EMBL: AC073395; AAG50971.1; -
KW Hypothetical protein.
SQ SEQUENCE 358 AA; 40239 MW; C1341019B1B63E CRC64;

Query Match 81.1%; Score 30; DB 10; Length 358;
Best Local Similarity 77.8%; Pred. No. 84;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OC NCBI_TaxId=3847;
RN [1]
RP SEQUENCE FROM N.A.
RA Karakaya C.H., Knap H.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF349572; AAK37555.1; -
DR InterPro: IPR001440; TRP.
SQ SEQUENCE 359 AA; 40243 MW; 306B46BC6E39EB2 CRC64;

Query Match 81.1%; Score 30; DB 10; Length 359;
Best Local Similarity 66.7%; Pred. No. 84;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OC 1 AVNAXSLF 9
DB 42 AYDALSLF 50

RESULT 14
OC9C780 PRELIMINARY; PRT; 371 AA.
ID 09C780
AC 09C780
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE Hypothetical protein.
GN AT3G1150/F11B9.109.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotsid II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=1130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansgorge W., Unsel'd M.,
RA Faltmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
RA Delbeny M., Boutry M., Grivell L.A., Maché R., Fulgomech P.,
RA De Simone V., Choisme N., Artiguenave F., Robert C., Broillet P.,
RA Winkler P., Catolico L., Weissendbach U., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmbach E., Drzonek H., Ertle H., Jordan N., Bangert S.,
RA Wiedemann R., Kranz H., Voess H., Holland R., Brandt P., Nakamura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppe S., Simonati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,
RA Reichelt J., Scharfe M., Schoen O., Barges M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Otenwaelder B., Duchemin D.,
RA Cooke R., Landie M., Berger-Llauro C., Fumelle B., Masny D.,
RA de Han M., Maarse A.C., Alcaraz J.-P., Cortet A., Casacuberta E.,
RA Monfort A., Argirou A., Flores M., Ligouri R., Vitale D.,
RA Mammucip G., Haase D., Schopf H., Radd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.K., Kaul S., Town C.D., Koo H.L., Talon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Utechtack T., Fujii C.Y., Shen T.P.,
RA Creasy T.H., Haas B., Walts R., Wu D., Peterson U., Van Aken S.,
RA Pai G., Miltscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nieman W.C., Salberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Ideasa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shino S., Tabata S., Wada T.,
RA Metarabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
thaliana.";
RL Nature 408:820-822(2000).
DR EMBL: AC073395; AAG50971.1; -
KW Hypothetical protein.
SQ SEQUENCE 371 AA; 41708 MW; EAE35CC1DEB6321B CRC64;

Query Match 81.1%; Score 30; DB 10; Length 371;
Best Local Similarity 77.8%; Pred. No. 87;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```
RT toxins."
RL Insect Biochem. Mol. Biol. 30:215-224(2000).
DR EMBL; AF034484; AAC36147.1; -.
DR MEROPS; M01.013; -.
DR InterPro; IPR001930; Ala_peptase.
DR InterPro; IPR006025; Zn_MTPeptase.
DR Pfam; PF01433; Peptidase_M1; 1.
DR PRINTS; PR00756; ALADIPPTASE.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 1016 AA; 115070 MW; C0D82D1FC582B33A CRC64;

Query Match      81.1%; Score 30; DB 5; Length 1016;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVNXXSLF 9
   |||||
DB 152 AVNATSAVY 160
```

Search completed: January 29, 2004, 14:56:48
Job time : 28.8676 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:52:19 ; Search time 11.1176 Seconds
(without alignments)
34.252 Million cell updates/sec

Title: US-10-032-950-4
Perfect score: 37
Sequence: 1 AVNAXSLP 9

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/ECTUS.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	94.6	9	US-09-339-511-4	Sequence 4, Appli
2	31	83.8	9	US-09-339-511-1	Sequence 1, Appli
3	30	81.1	9	US-08-861-476C-3	Sequence 3, Appli
4	30	81.1	47	US-08-861-476C-6	Sequence 6, Appli
5	29	78.4	426	US-09-198-452A-760	Sequence 760, App
6	29	78.4	1054	US-07-596-467-4	Sequence 4, Appli
7	29	78.4	1054	US-07-934-374-4	Sequence 4, Appli
8	29	78.4	1054	US-07-783-861C-2	Sequence 2, Appli
9	28	75.7	9	US-09-339-511-5	Sequence 5, Appli
10	28	75.7	9	US-09-339-511-6	Sequence 6, Appli
11	27	73.0	9	US-09-339-511-2	Sequence 2, Appli
12	27	73.0	518	US-09-625-972-23	Sequence 23, Appli
13	27	73.0	762	US-09-252-991A-27566	Sequence 27566, A
14	27	73.0	1720	US-08-477-451-12	Sequence 12, Appli
15	26	70.3	9	US-09-339-511-7	Sequence 7, Appli
16	26	70.3	223	US-09-107-532A-3574	Sequence 3674, Ap
17	26	70.3	229	US-09-134-001C-4241	Sequence 4241, Ap
18	26	70.3	221	US-09-252-991A-20970	Sequence 20970, A
19	26	70.3	338	US-07-809-457A-16	Sequence 16, Appli
20	26	70.3	338	US-08-553-943-16	Sequence 16, Appli
21	26	70.3	338	PCT-US91-09437-16	Sequence 16, Appli
22	26	70.3	364	US-08-860-820-2	Sequence 2, Appli
23	26	70.3	390	US-09-419-459-4	Sequence 4, Appli
24	26	70.3	394	US-09-419-459-10	Sequence 10, Appli
25	26	70.3	395	US-09-328-352-5819	Sequence 5819, Ap
26	26	70.3	582	US-09-419-459-2	Sequence 2, Appli
27	26	70.3	1471	US-08-755-587-188	Sequence 188, App

28	25	67.6	35	US-09-205-258-436	Sequence 436, App
29	25	67.6	103	US-09-240-274-42	Sequence 42, Appli
30	25	67.6	304	US-08-252-991A-18282	Sequence 18282, A
31	25	67.6	333	US-08-118-270-40	Sequence 40, Appli
32	25	67.6	333	PCT-US93-08528-40	Sequence 40, Appli
33	25	67.6	335	US-09-106-872A-17	Sequence 17, Appli
34	25	67.6	347	US-09-215-418-5	Sequence 5, Appli
35	25	67.6	352	US-09-198-452A-702	Sequence 702, App
36	25	67.6	364	US-08-483-101-5	Sequence 5, Appli
37	25	67.6	373	US-08-071-035-116	Sequence 116, App
38	25	67.6	387	US-08-484-575A-7	Sequence 7, Appli
39	25	67.6	387	US-08-477-459-7	Sequence 7, Appli
40	25	67.6	387	US-08-479-869-7	Sequence 7, Appli
41	25	67.6	387	US-08-486-414-7	Sequence 7, Appli
42	25	67.6	387	PCT-US94-01826A-7	Sequence 7, Appli
43	25	67.6	387	PCT-US94-02252A-7	Sequence 7, Appli
44	25	67.6	406	US-09-071-035-114	Sequence 114, App
45	25	67.6	410	US-07-955-905A-26	Sequence 26, Appli

ALIGNMENTS

```

RESULT 1
US-09-339-511-4
; Sequence 4, Application US/09339511
; Patent No. 6337385
; GENERAL INFORMATION:
; APPLICANT: Muir, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. 6337385ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/09/339, 511
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-4
Query Match          94.6% Score 35; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches          9; Conservative          0; Mismatches          0; Indels          0; Gaps          0;
QY          1 AVNAXSLP 9
Db          1 AVNAXSLP 9
RESULT 2
US-09-339-511-1
; Sequence 1, Application US/09339511
; Patent No. 6337385
; GENERAL INFORMATION:
; APPLICANT: Muir, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. 6337385ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong

```

TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/09/339,511
CURRENT FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-1

Query Match 83.8%; Score 31; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VNAXSSLP 9
DB 2 VNAXSSLP 9

RESULT 3
US-08-861-476C-3
Sequence 3, Application US/08861476C
Patent No. 6447786
GENERAL INFORMATION:
APPLICANT: New York University Medical Center
TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
FILE REFERENCE: 63753/7
CURRENT APPLICATION NUMBER: US/08/861,476C
CURRENT FILING DATE: 1997-05-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Version 3.0
SEQ ID NO 3
LENGTH: 9
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-08-861-476C-3

Query Match 81.1%; Score 30; DB 4; Length 9;
Best Local Similarity 87.5%; Pred. No. 2.5e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSSLP 9
DB 2 VNAXSSLP 9

RESULT 4
US-08-861-476C-6
Sequence 6, Application US/08861476C
Patent No. 6447786
GENERAL INFORMATION:
APPLICANT: New York University Medical Center
TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
FILE REFERENCE: 63753/7
CURRENT APPLICATION NUMBER: US/08/861,476C
CURRENT FILING DATE: 1997-05-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Version 3.0
SEQ ID NO 6
LENGTH: 47
TYPE: PRT
ORGANISM: Staphylococcus aureus

US-08-861-476C-6

Query Match 81.1%; Score 30; DB 4; Length 47;
Best Local Similarity 87.5%; Pred. No. 6.3;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSSLP 9
DB 25 VNAXSSLP 32

RESULT 5
US-09-198-452A-760
Sequence 760, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Griffiths, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment and uses thereof, in particular for the diagnosis, prevention and treatment of infection
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 760
LENGTH: 426
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-760

Query Match 78.4%; Score 29; DB 4; Length 426;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNALSLP 9
DB 116 AVNALSLP 124

RESULT 6
US-07-596-467-4
Sequence 4, Application US/07596467
Patent No. 5306862
GENERAL INFORMATION:
APPLICANT: Chappell, J.
APPLICANT: Saunders, Court A.
APPLICANT: Wolf, Fred R.
TITLE OF INVENTION: Method and Composition for Increasing Sterol Accumulation in Higher Plants
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dressler, Goldsmith, Shore, Suter & Milnamow
STREET: 180 N. Stetson St.
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/596,467
CLASSIFICATION: 800
FILING DATE: 19901012
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312 616-5400
TELEFAX: 312 616-5460
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1054 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-596-467-4

Query Match 78.4%; Score 29; DB 1; Length 1054;
 Best Local Similarity 75.0%; Pred. No. 3.2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSSLF 9
 ||: ||||
 Db 226 VNAXSSLF 233

RESULT 7
 US-07-934-374-4
 ; Sequence 4, Application US/07934374
 ; Patent No. 5349126
 ; GENERAL INFORMATION:

APPLICANT: Chappell, J.
 APPLICANT: Saunders, Court A.
 APPLICANT: Cusellar, R.
 TITLE OF INVENTION: Process and Composition for Increasing
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dresler, Goldsmith, Shore, Sutter & Milnamow
 STREET: 180 N. Stearns St.
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60601

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/934,374
 FILING DATE: 19920814
 CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:
 NAME: Gansson, Edward P.
 REGISTRATION NUMBER: 29,381
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312 616-5460
 TELEFAX: 312 616-5460

INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1054 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-934-374-4

Query Match 78.4%; Score 29; DB 1; Length 1054;
 Best Local Similarity 75.0%; Pred. No. 3.2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSSLF 9
 ||: ||||
 Db 226 VNAXSSLF 233

RESULT 8
 US-07-783-861C-2
 ; Sequence 2, Application US/07783861C
 ; Patent No. 5460949
 ; GENERAL INFORMATION:
 APPLICANT: Saunders, Court A.

APPLICANT: Wolf, Fred R.
 APPLICANT: Mukharji, Indrani
 TITLE OF INVENTION: A Method and Composition for Increasing
 TITLE OF INVENTION: the Accumulation of Squalene and Specific Steroids in
 TITLE OF INVENTION: Yeast
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Amoco Corp., Patents and Licensing Dept.
 STREET: 200 East Randolph St.
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60680-0703

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/783,861C
 FILING DATE: 19911028
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/613,380
 FILING DATE: 15-NOV-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Galloway, No. 5460949v11 B.
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312 856-4972

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1054 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-783-861C-2

Query Match 78.4%; Score 29; DB 1; Length 1054;
 Best Local Similarity 75.0%; Pred. No. 3.2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSSLF 9
 ||: ||||
 Db 226 VNAXSSLF 233

RESULT 9
 US-09-339-511-5
 ; Sequence 5, Application US/09339511
 ; Patent No. 6337385
 ; GENERAL INFORMATION:

APPLICANT: Muir, Tom
 APPLICANT: Mayville, Patricia
 APPLICANT: No. 6337385ick, Richard P.
 APPLICANT: Beavis, Ronald
 APPLICANT: Ji, Guangyong

TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
 TITLE OF INVENTION: INTERFERENCE
 FILE REFERENCE: 600-1-231N
 CURRENT APPLICATION NUMBER: US/09/339,511
 CURRENT FILING DATE: 1999-06-24
 PRIOR APPLICATION NUMBER: 60/090,402
 PRIOR FILING DATE: 1998-06-24
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 5

LENGTH: 9
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE: Description of Artificial Sequence: Synthetic
 OTHER INFORMATION: peptide

NAME/KEY: VARIANT
 LOCATION: (5)
 OTHER INFORMATION: Xaa represents any amino acid at this position.
 US-09-339-511-5

Query Match 75.7%; Score 28; DB 4; Length 9;
 Best Local Similarity 87.5%; Pred. No. 2.5e+05;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VNAXSLF 9
 Db 2 VNAXSLF 9

RESULT 10
 US-09-339-511-6
 Sequence 6, Application US/09339511
 Patent No. 6337385

GENERAL INFORMATION:
 APPLICANT: Mult, Tom
 APPLICANT: Mayville, Patricia
 APPLICANT: No. 6337385ick, Richard P.
 APPLICANT: Beavis, Ronald
 APPLICANT: Ji, Guangyong
 TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
 TITLE OF INVENTION: INTERFERENCE
 FILE REFERENCE: 600-1-221N
 CURRENT APPLICATION NUMBER: US/09/339,511
 CURRENT FILING DATE: 1999-06-24
 PRIOR APPLICATION NUMBER: 60/090,402
 PRIOR FILING DATE: 1998-06-24
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 6
 LENGTH: 9
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 OTHER INFORMATION: peptide
 NAME/KEY: VARIANT
 LOCATION: (5)
 OTHER INFORMATION: Xaa represents any amino acid at this position.
 US-09-339-511-6

Query Match 75.7%; Score 28; DB 4; Length 9;
 Best Local Similarity 87.5%; Pred. No. 2.5e+05;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VNAXSLF 9
 Db 2 VNAXSLF 9

RESULT 11
 US-09-339-511-2
 Sequence 2, Application US/09339511
 Patent No. 6337385
 GENERAL INFORMATION:
 APPLICANT: Mult, Tom
 APPLICANT: Mayville, Patricia
 APPLICANT: No. 6337385ick, Richard P.
 APPLICANT: Beavis, Ronald
 APPLICANT: Ji, Guangyong
 TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
 TITLE OF INVENTION: INTERFERENCE
 FILE REFERENCE: 600-1-221N
 CURRENT APPLICATION NUMBER: US/09/339,511
 CURRENT FILING DATE: 1999-06-24
 PRIOR APPLICATION NUMBER: 60/090,402
 PRIOR FILING DATE: 1998-06-24
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 2
 LENGTH: 9
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 OTHER INFORMATION: peptide
 NAME/KEY: VARIANT
 LOCATION: (5)
 OTHER INFORMATION: Xaa represents any amino acid at this position.
 US-09-339-511-2

Query Match 73.0%; Score 27; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 NAXSLF 9
 Db 3 NAXSLF 9

RESULT 12
 US-09-625-972-23
 Sequence 23, Application US/09625972
 Patent No. 6566513
 GENERAL INFORMATION:
 APPLICANT: GIBERTIER, Iutz Gerhard
 APPLICANT: HAUSER, Hans Peter
 APPLICANT: DOMINGO DELORO, Yvette Beatrice
 APPLICANT: ZEKENG, Leopold
 APPLICANT: KAPITE, Lazare
 TITLE OF INVENTION: LENTIVIRUS FROM THE GROUP OF IMMUNODEFICIENCY VIRUSES OF DRILL
 TITLE OF INVENTION: (MANDRILLUS LEUCOPHAERUS) AND THEIR USE
 FILE REFERENCE: 067595/0106
 CURRENT APPLICATION NUMBER: US/09/625,972
 CURRENT FILING DATE: 2000-07-29
 PRIOR APPLICATION NUMBER: DE 199 36 003.0
 PRIOR FILING DATE: 1999-08-03
 NUMBER OF SEQ ID NOS: 57
 SOFTWARE: Patentin version 3.0
 SEQ ID NO 23
 LENGTH: 518
 TYPE: PRT
 ORGANISM: SV - Viral
 US-09-625-972-23

Query Match 73.0%; Score 27; DB 4; Length 518;
 Best Local Similarity 66.7%; Pred. No. 3.9e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 AVNAXSLF 9
 Db 504 AVNAXSLF 512

RESULT 13
 US-09-252-991A-27566
 Sequence 27566, Application US/09252991A
 Patent No. 6551795
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfeld et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 27566
 LENGTH: 762

TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27586

Query Match
Best Local Similarity 73.0%; Score 27; DB 4; Length 762;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
DB 596 AVNAXSLF 604

RESULT 14
US-08-477-451-12
Sequence 12, Application US/08477451
Patent No. 5928865
GENERAL INFORMATION:
APPLICANT: Covacci, Antonello
TITLE OF INVENTION: Helicobacter Pylori CagI Region
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,451
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McClung, Barbara G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: 0335.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2708
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1720 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-477-451-12

Query Match
Best Local Similarity 73.0%; Score 27; DB 2; Length 1720;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VNAXSLF 9
DB 1664 INATQSLF 1671

RESULT 15
US-09-339-511-7
Sequence 7, Application US/09339511
Patent No. 6337385
GENERAL INFORMATION:
APPLICANT: Mair, Tom
APPLICANT: Mayville, Patricia
APPLICANT: No. 6337385ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: Ji, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL

TITLE OF INVENTION: INTERFERENCE
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/09/339,511
CURRENT FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: 60/050,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 7
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-7

Query Match
Best Local Similarity 70.3%; Score 26; DB 4; Length 9;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSLF 9
DB 2 VNAXSSAP 9

Search completed: January 29, 2004, 15:03:21
Job time: 11.1176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:47:39 ; Search time 25.5441 Seconds

(without alignments)
73.223 Million cell updates/sec

Title: US-10-032-950-4

Sequence: 37
1 AVNAXSLF 9Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*

10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*

11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/prodata/1/pubpaa/US09C_NEW_PUB.pep:*

13: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*

14: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*

15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/prodata/1/pubpaa/US10C_NEW_PUB.pep:*

17: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep:*

18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	94.6	9	US-10-032-950-4	Sequence 4, Appli
2	34	91.9	633	US-10-369-493-6225	Sequence 6225, Ap
3	31	83.8	9	US-10-032-950-1	Sequence 1, Appli
4	30	81.1	9	US-10-201-444-3	Sequence 6, Appli
5	30	81.1	47	US-10-201-444-6	Sequence 3, Appli
6	30	81.1	572	US-09-919-781-2	Sequence 2, Appli
7	30	81.1	572	US-10-154-419-13	Sequence 13, Appli
8	29	78.4	426	US-10-289-762-760	Sequence 760, App
9	29	78.4	605	US-10-100-303A-110	Sequence 110, App
10	29	78.4	1054	US-10-369-493-1449	Sequence 1449, Ap
11	28	75.7	9	US-10-032-950-5	Sequence 5, Appli
12	28	75.7	9	US-10-032-950-6	Sequence 6, Appli
13	28	75.7	78	US-10-078-929-136	Sequence 136, App
14	28	75.7	248	US-10-369-493-14264	Sequence 14264, A
15	28	75.7	248	US-10-369-493-14770	Sequence 14770, A

16	28	75.7	257	12	US-10-369-493-11482	Sequence 11482, A
17	28	75.7	441	12	US-10-369-493-2265	Sequence 2265, Ap
18	28	75.7	526	12	US-10-369-493-3924	Sequence 3924, Ap
19	28	75.7	561	14	US-10-078-929-138	Sequence 138, App
20	27	73.0	9	13	US-10-032-950-2	Sequence 2, Appli
21	27	73.0	61	13	US-10-029-386-27613	Sequence 27613, A
22	27	73.0	118	12	US-10-173-551-14	Sequence 14, Appl
23	27	73.0	159	12	US-10-407-079-37	Sequence 37, Appl
24	27	73.0	379	14	US-10-047-460-12	Sequence 12, Appl
25	27	73.0	492	12	US-10-369-493-31	Sequence 31, Appl
26	27	73.0	518	12	US-10-364-360-23	Sequence 23, Appl
27	27	73.0	726	12	US-10-369-493-6788	Sequence 6788, Ap
28	27	73.0	765	14	US-10-044-303-1	Sequence 1, Appli
29	27	73.0	1016	14	US-10-044-303-2	Sequence 2, Appli
30	27	73.0	1289	10	US-09-932-145-11	Sequence 11, Appl
31	27	73.0	1289	12	US-10-395-812-11	Sequence 11, Appl
32	26	70.3	9	13	US-10-032-950-7	Sequence 7, Appli
33	26	70.3	300	10	US-09-738-626-4143	Sequence 4143, Ap
34	26	70.3	356	15	US-10-156-661-9876	Sequence 9876, Ap
35	26	70.3	379	12	US-10-369-493-6990	Sequence 6990, Ap
36	26	70.3	440	12	US-10-310-154-733	Sequence 733, App
37	26	70.3	542	12	US-10-369-493-218	Sequence 218, App
38	26	70.3	727	10	US-09-815-923-10	Sequence 10, Appl
39	26	70.3	908	12	US-10-085-959-255	Sequence 255, App
40	26	70.3	1087	12	US-10-369-493-21095	Sequence 21095, A
41	26	70.3	1511	12	US-10-369-493-22496	Sequence 22496, A
42	26	70.3	1684	12	US-10-328-075-877	Sequence 877, App
43	25	67.6	35	12	US-09-933-767-436	Sequence 436, App
44	25	67.6	35	15	US-10-023-482-436	Sequence 436, App
45	25	67.6	55	10	US-09-764-877-2024	Sequence 2024, Ap

ALIGNMENTS

RESULT 1

US-10-032-950-4

Sequence 4, Application US/10032950

Publication No. US20020077453A1

GENERAL INFORMATION:

APPLICANT: Malt, Tom

APPLICANT: Mayville, Patricia

APPLICANT: No. US20020077453A1ick, Richard P.

APPLICANT: Beavis, Ronald

APPLICANT: Ji, Guangyong

TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL

TITLE OF INVENTION: INTERFERENCE

FILE REFERENCE: 600-1-231N

CURRENT APPLICATION NUMBER: US/10/032, 950

CURRENT FILING DATE: 2001-12-27

PRIOR APPLICATION NUMBER: 60/090,402

PRIOR FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 8

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 4

LENGTH: 9

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

NAME/KEY: VARIANT

LOCATION: (5)

OTHER INFORMATION: Xaa represents any amino acid at this position.

US-10-032-950-4

Query Match 94.6%; Score 35; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
Db 1 AVNAXSLF 9

RESULT 2
US-10-369-493-6225
; Sequence 6225, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianning
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6225
; LENGTH: 633
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6225

Query Match 91.9%; Score 34; DB 12; Length 633;
Best Local Similarity 77.8%; Pred. No. 21;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VNAXSSLF 9
DB 327 VNAXSSLF 335

RESULT 3
US-10-032-950-1
; Sequence 1, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Malt, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-1

Query Match 93.8%; Score 31; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VNAXSSLF 9
DB 2 VNAXSSLF 9

RESULT 4
US-10-201-444-3
; Sequence 3, Application US/10201444
; Publication No. US20030078378A1
; GENERAL INFORMATION:
; APPLICANT: New York University Medical Center
; TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
; FILE REFERENCE: 63753/7
; CURRENT APPLICATION NUMBER: US/10/201,444
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/08/861,476
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentn version 3.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-201-444-3

Query Match 81.1%; Score 30; DB 15; Length 9;
Best Local Similarity 87.5%; Pred. No. 7e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSSLF 9
DB 2 VNAXSSLF 9

RESULT 5
US-10-201-444-6
; Sequence 6, Application US/10201444
; Publication No. US20030078378A1
; GENERAL INFORMATION:
; APPLICANT: New York University Medical Center
; TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
; FILE REFERENCE: 63753/7
; CURRENT APPLICATION NUMBER: US/10/201,444
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/08/861,476
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentn version 3.0
; SEQ ID NO 6
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-201-444-6

Query Match 81.1%; Score 30; DB 15; Length 47;
Best Local Similarity 87.5%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAXSSLF 9
DB 25 VNAXSSLF 32

RESULT 6
US-09-919-781-2
; Sequence 2, Application US/09919781
; Patent No. US20020123094A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: 57250, A NOVEL HUMAN SUGAR TRANSPORTER
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: NMT-181
; CURRENT APPLICATION NUMBER: US/09/919,781
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: USSN 60/221,769
; PRIOR FILING DATE: 2000-07-31

NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 572
TYPE: PRT
ORGANISM: Homo sapiens
US-09-919-781-2

Query Match 81.1%; Score 30; DB 10; Length 572;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
DB 453 AVNAXSLF 461

RESULT 7
US-10-154-419-13

Sequence 13, Application US/10154419
Publication No. US2003013675A1
GENERAL INFORMATION:
APPLICANT: Curtis, Rory A.J.
APPLICANT: Gluckman, Maria Alexandra
APPLICANT: Meyers, Rachel E.
TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8099, 46455, 54414, 53763,
TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FBH56295FL, 57255,
TITLE OF INVENTION: AND 57255alt MOLECULES AND USES THEREFOR
FILE REFERENCE: MNI-249
CURRENT APPLICATION NUMBER: US/10/154,419
CURRENT FILING DATE: 2002-05-22
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 99
SOFTWARE: FastSeq Version 4.0
SEQ ID NO 13
LENGTH: 572
TYPE: PRT
ORGANISM: Homo sapiens
US-10-154-419-13

Query Match 81.1%; Score 30; DB 12; Length 572;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
DB 453 AVNAXSLF 461

RESULT 8
US-10-289-762-760
Sequence 760, Application US/10289762
Publication No. US20040006218A1
GENERAL INFORMATION:
APPLICANT: Griffiths, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/10/289,762
CURRENT FILING DATE: 2003-03-27
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 760
LENGTH: 426
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-10-289-762-760

Query Match 78.4%; Score 29; DB 12; Length 426;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
DB 116 AVNAXSLF 124

RESULT 9
US-10-100-303A-110
Sequence 110, Application US/10100303A
Publication No. US2003020280A1
GENERAL INFORMATION:
APPLICANT: Caplan, et al.
TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
TITLE OF INVENTION: to Allergy
FILE REFERENCE: 2002834-0166
CURRENT APPLICATION NUMBER: US/10/100,303A
CURRENT FILING DATE: 2002-03-18
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 110
LENGTH: 605
TYPE: PRT
ORGANISM: Glycine Max
US-10-100-303A-110

Query Match 78.4%; Score 29; DB 12; Length 605;
Best Local Similarity 75.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAWSLNF 9
DB 517 VNAWSLNF 524

RESULT 10
US-10-369-493-1849
Sequence 1849, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 1849
LENGTH: 1054
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1849

Query Match 78.4%; Score 29; DB 12; Length 1054;
Best Local Similarity 75.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNAWSLNF 9
DB 226 VNAWSLNF 233

RESULT 11
US-10-032-950-5
Sequence 5, Application US/10032950
Publication No. US2002007453A1
GENERAL INFORMATION:
APPLICANT: Muir, Tom
APPLICANT: Mayville, Patricia


```

; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-5

```

```

Query Match          75.7%; Score 28; DB 13; Length 9;
Best Local Similarity 87.5%; Pred. No. 7e+05; 0; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 0;

```

```

QY      2 VNAXSSLF 9
       |||||
Db      2 VNAXSLF 9

```

```

RESULT 12
US-10-032-950-6
; Sequence 6, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mullr, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-6

```

```

Query Match          75.7%; Score 28; DB 13; Length 9;
Best Local Similarity 87.5%; Pred. No. 7e+05; 0; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 0;

```

```

QY      2 VNAXSSLF 9
       |||||
Db      2 VNAXSLF 9

```

```

RESULT 13
US-10-078-929-136
; Sequence 136, Application US/10078929
; Publication No. US20020152497A1
; GENERAL INFORMATION:
; APPLICANT: Rafalski, Antoni
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Sakai, Hajime
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Odell, Joan T.
; APPLICANT: Meyers, Blake
; APPLICANT: Thorpe, Catherine
; APPLICANT: Meng, Zude
; TITLE OF INVENTION: Nucleic Acid Fragments Encoding Proteins Involved in
; FILE REFERENCE: B1357 US NA
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 09/566,394
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 60/133038
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133042
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133427
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133437
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133428
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133438
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133436
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/137667
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 136
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Trifolium aestivum
US-10-078-929-136

```

```

Query Match          75.7%; Score 28; DB 14; Length 78;
Best Local Similarity 75.0%; Pred. No. 50;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 AVNAXSSL 8
       |||||
Db      9 AINAXSSL 16

```

```

RESULT 14
US-10-369-493-14264
; Sequence 14264, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10152052/B
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14264

```

; LENGTH: 248
 ; TYPE: PRT
 ; ORGANISM: Agrobacterium tumefaciens
 US-10-369-493-14264

Query Match 75.7%; Score 28; DB 12; Length 248;
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
 |||||:
 Db 113 AVNAXSAFF 121

RESULT 15
 US-10-369-493-14770
 ; Sequence 14770; Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ. ID NOS: 47374
 ; SEQ. ID NO 14770
 ; LENGTH: 248
 ; TYPE: PRT
 ; ORGANISM: Agrobacterium tumefaciens
 US-10-369-493-14770

Query Match 75.7%; Score 28; DB 12; Length 248;
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVNAXSLF 9
 |||||:
 Db 113 AVNAXSAFF 121

Search completed: January 29, 2004, 15:01:45
 Job time : 25.5441 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:44:54 : Search time 35.0735 Seconds
(without alignments)
40.730 Million cell updates/sec

Title: US-10-032-950-5
Perfect score: 39
Sequence: 1 GVAXASLRF 9

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	94.9	9	21	AAV67855
2	37	94.9	9	23	ABP53544
3	37	94.9	9	23	AAV50903
4	35	89.7	9	21	AAV67859
5	35	89.7	9	21	AAV67860
6	35	89.7	9	23	AAV50907
7	35	89.7	9	23	AAV51002
8	35	89.7	9	23	AAV51003
9	35	89.7	9	23	AAV51004

10	35	89.7	9	23	ABB07160
11	34	87.2	9	21	AAV67851
12	34	87.2	9	21	AAV67861
13	34	87.2	9	23	ABP53540
14	34	87.2	9	23	AAV50898
15	34	87.2	9	23	AAV50908
16	34	87.2	9	23	AAV51005
17	34	87.2	9	23	AAV67161
18	33	84.6	9	19	AAV38323
19	33	84.6	9	23	AAV51001
20	33	84.6	9	23	ABB07159
21	33	84.6	9	24	ABB84631
22	33	84.6	47	21	ABB84634
23	31	79.5	9	21	AAV67856
24	31	79.5	9	23	ABP53545
25	31	79.5	9	23	AAV50904
26	30	76.9	9	21	AAV67852
27	30	76.9	9	23	ABP53541
28	30	76.9	9	23	AAV50900
29	30	76.9	147	21	ABA11086
30	30	76.9	147	23	ABP34694
31	30	76.9	552	22	ABG13082
32	29	74.4	9	21	AAV67857
33	29	74.4	9	23	ABP53546
34	29	74.4	9	23	AAV50905
35	29	74.4	198	22	ABG18065
36	29	74.4	332	23	ABB05389
37	29	74.4	397	20	AAV37161
38	29	74.4	488	16	AAV02271
39	29	74.4	488	22	ABV56384
40	29	74.4	488	24	ABP81905
41	29	74.4	490	22	ABG26292
42	29	74.4	605	23	ABP94351
43	29	74.4	612	22	ABV92352
44	29	74.4	747	20	AAV41758
45	29	74.4	747	21	AAV44314

ALIGNMENTS

RESULT 1	
AAV67855	
ID	AAV67855 standard; peptide; 9 AA.
AC	AAV67855;
XX	
DT	25-APR-2000 (first entry)
XX	
DE	S. aureus peptide #5 used for bacterial interference.
XX	
KW	Staphylococcus aureus infection; cyclic peptide; AgrD; agr response;
XX	
KW	virulence factor; treatment.
XX	
OS	Staphylococcus aureus.
XX	
XX	
FT	Key Location/Qualifiers
FT	Modified-site 1
FT	/note="N-terminal residue forms bond with C-terminal residue to form a cyclic peptide"
FT	Misc-difference 5
FT	/label= Unknown
XX	
PN	WO9967286-A2.
XX	
PD	29-DEC-1999.
XX	
PF	24-JUN-1999; 95MO-US14562.
XX	
PR	24-JUN-1998; 98US-0103438.
XX	
XX	
PA	(VYRQ) UNIV ROCKEFELLER.
PA	(VYNY) UNIV NEW YORK STATE.

Peptide-mediated q
S. aureus peptide
Staphylococcus aur
Cyclic peptide SEQ
AgrD-autoinducing
Protected peptide
AgrD2 lactam cycli
Transcription inh
AgrD2 chiolacton
Peptide-mediated q
S. aureus RN6607 A
S. aureus RN6607 A
S. aureus peptide
Cyclic peptide SEQ
AgrD-autoinducing
S. aureus peptide
Cyclic peptide SEQ
AgrD-autoinducing
Human OREFX ORF850
Human ORF367, prot
Novel human diagn
S. aureus peptide
Cyclic peptide SEQ
AgrD-autoinducing
Novel human diagn
Human 22109 protei
Amino acid sequenc
Human prostaglandi
Non-endogenous hum
Human prostaglandi
Novel human diagn
Chlamydia trachoma
Human protein sequ
Human PRO102 proc
Human PRO102 (UNQ

```

XX  Muir TW, Mayville P, Novick RP, Ji G, Beavis R,
PI  WPI; 2000-147202/13.
XX  New cyclic peptides for treating infections with Staphylococcus aureus
PT  -
XX  Claim 9; Page 26; 37pp; English.
XX  This sequence represents a cyclic peptide derived from the Staphylococcus
CC  aureus AgrD peptide. The invention relates to AgrD derived peptides,
CC  a composition containing a peptide and a carrier, and a method for the
CC  production of the cyclic peptides. The peptide inhibits the agr response,
CC  which is normally associated with the release of virulence factors of
CC  Staphylococcus aureus. An AgrD peptide is produced by S. aureus that
CC  activates the agr response in strains of a single group, but interferes
CC  with this response in strains of different groups. The peptides and
CC  composition containing them can be used to treat infections by S. aureus.
XX  Sequence 9 AA;
SQ
XX  Query Match 94.9%; Score 37; DB 21; Length 9;
XX  Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY  1 GVNAXASLRF 9
DB  1 GVNAXASLRF 9
XX  RESULT 2
XX  ABP53544 ID ABP53544 standard; peptide; 9 AA.
XX  AC ABP53544;
XX  DT 13-DEC-2002 (first entry)
XX  DE Cyclic peptide SEQ ID NO:5.
XX  KM Cyclic peptide; Staphylococcus aureus; infection; antibacterial;
XX  agr response inhibitor.
XX  OS Synthetic.
XX  FH Key Location/Qualifiers
XX  FT Misc-difference 5 /note="any amino acid"
XX  PN US2002077453-A1.
XX  PD 20-JUN-2002.
XX  PF 27-DEC-2001; 2001US-0032950.
XX  PR 24-JUN-1998; 98US-090402P.
XX  PR 24-JUN-1999; 99US-0339511.
XX  PA (MUIR/) MUIR T W.
XX  PA (MAYV/) MAYVILLE P.
XX  PA (NOVI/) NOVICK R P.
XX  PA (BEAV/) BEAVIS R.
XX  PA (JIGG/) JI G.
XX  PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX  DR WPI; 2002-681366/73.
XX  PT New cyclic peptides, useful for treating Staphylococcus aureus
XX  infections -
XX  PS Claim 9; Page 10; 18pp; English.

```

```

XX  ABP53540 to ABP53547 represent cyclic peptides (I) from the present
CC  invention. The present invention also describes a method for treating
CC  Staphylococcus aureus infection comprising the administration of a
CC  composition comprising (I). (I) has antibacterial activity, and can be
CC  used as an agr gene response inhibitor. The peptides are useful for
CC  treating S. aureus infections.
XX  Sequence 9 AA;
SQ
XX  Query Match 94.9%; Score 37; DB 23; Length 9;
XX  Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY  1 GVNAXASLRF 9
DB  1 GVNAXASLRF 9
XX  RESULT 3
XX  AAM50903 ID AAM50903 standard; peptide; 9 AA.
XX  AC AAM50903;
XX  DT 08-MAY-2002 (first entry)
XX  DE AgrD-autoinducing cyclic peptide, inhibitor of agr response.
XX  KM Staphylococcus aureus; AgrD; agr response; inhibitor; antibiotic;
XX  antibacterial; infection; therapy; cyclic.
XX  OS Synthetic.
XX  FH Key Location/Qualifiers
XX  FT Misc-difference 5 /note="any amino acid"
XX  FT Misc-difference 5 /note="note linked to residue 9 to form cyclic
XX  peptide"
XX  FT Misc-difference 9 /note="note linked to residue 5 to form cyclic
XX  peptide"
XX  PN US6337385-B1.
XX  PD 08-JAN-2002.
XX  PF 24-JUN-1999; 99US-0339511.
XX  PR 24-JUN-1998; 98US-090402P.
XX  PA (UYRO ) UNIV ROCKEFELLER.
XX  PA (UYNY ) UNIV NEW YORK STATE.
XX  PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX  DR WPI; 2002-170774/22.
XX  PT Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
XX  interference and for treating Staphylococcus aureus infection in a
XX  subject -
XX  PS Claim 7; Column 19; 18pp; English.
XX  The present sequence is that of a novel synthetic cyclic peptide
CC  of the invention that is capable of inhibiting the agr response of
CC  Staphylococcus aureus. It is an AgrD-autoinducing peptide, where
CC  AgrD is a secreted agr-encoded peptide and where the agr locus
CC  controls the synthesis of virulence factor and other extracellular
CC  proteins responsible for pathogenicity in S. aureus. Preferred
CC  peptides may have the sequence NH2-X(n)-Z-X(Y)-COOH, with a
CC  cyclic bond between the Z residue and COOH other than a thioester

```

CC bond, where X is an amino acid, an amino acid analogue, a
 CC peptidomimetic or non-amide isostere, Z is a synthetic or a
 CC diosynthetic amino acid, n is 0-10 and Y is 1-10. The cyclic bond
 CC is especially a lactam or lactone bond. The thiololactone
 CC structure within native Agd peptides is required for activation
 CC of the agr response. Elimination of the thiol ester component of
 CC the cyclic ring structure can destroy agr response activating
 CC activity while preserving and enhancing inhibitory activity. A
 CC claimed method of preparing a cyclic peptide involves assembling
 CC a linear peptide chain on to a solid phase resin support; treating the
 CC deprotected the resulting protected assembled peptide; treating the
 CC deprotected peptide with neutral buffer for a time sufficient to
 CC form the cyclic peptide and cleave the peptide from the support;
 CC and recovering the cyclic peptide. The peptide is useful for
 CC bacterial interference, especially for the treatment of *S. aureus*
 CC infection.

SQ Sequence 9 AA;
 XX
 XX

Query Match 94.9%; Score 37; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GVNAXASLF 9
 DB 1 GVNAXASLF 9

RESULT 4
 AAY67859
 ID AAY67859 standard; peptide; 9 AA.
 XX
 XX
 AC AAY67859;
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE Staphylococcus aureus AgdII derived peptide sequence.
 XX
 XX Staphylococcus aureus infection; treatment; AgrD; agr response;
 KM virulence factor.
 KW
 OS Staphylococcus aureus.
 XX
 PN WO9967286-A2.
 XX
 PD 29-DEC-1999.
 XX
 PF 24-JUN-1999; 99WO-US14562.
 XX
 PR 24-JUN-1998; 98US-0103438.
 XX
 PA (UYRO) UNIV ROCKEFELLER.
 XX (UYNY) UNIV NEW YORK STATE.
 PI Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
 XX
 XX WPI; 2000-147202/13.
 XX
 DR
 PT New cyclic peptides for treating infections with Staphylococcus aureus
 XX
 PS Examples; Page 22; 37pp; English.
 XX
 CC This sequence represents the Staphylococcus aureus AgdII derived
 CC peptide. The invention relates to Agd derived cyclic peptides, a
 CC composition containing a peptide and a carrier, and a method for the
 CC production of the cyclic peptides. The peptide inhibits the agr response,
 CC which is normally associated with the release of virulence factors of
 CC Staphylococcus aureus. An Agd peptide is produced by *S. aureus* that
 CC activates the agr response in strains of a single group, but interferes
 CC with this response in strains of different groups. The peptides and
 CC composition containing them can be used to treat infections by *S. aureus*.

SQ Sequence 9 AA;
 XX
 XX

Query Match 89.7%; Score 35; DB 21; Length 9;
 Best Local Similarity 77.8%; Pred. No. 9.3e+05;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GVNAXASLF 9
 DB 1 GVNAXASLF 9

RESULT 5
 AAY67860
 ID AAY67860 standard; peptide; 9 AA.
 XX
 XX
 AC AAY67860;
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE Staphylococcus aureus AgdII derived peptide sequence.
 XX
 XX Staphylococcus aureus infection; AgrD; agr response; treatment;
 KM virulence factor.
 KW
 OS Staphylococcus aureus.
 XX
 PN WO9967286-A2.
 XX
 PD 29-DEC-1999.
 XX
 PF 24-JUN-1999; 99WO-US14562.
 XX
 PR 24-JUN-1998; 98US-0103438.
 XX
 PA (UYRO) UNIV ROCKEFELLER.
 XX (UYNY) UNIV NEW YORK STATE.
 PI Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
 XX
 XX WPI; 2000-147202/13.
 XX
 DR
 PT New cyclic peptides for treating infections with Staphylococcus aureus
 XX
 PS Examples; Page 22; 37pp; English.
 XX
 CC This sequence represents the Staphylococcus aureus AgdII derived
 CC peptide. The invention relates to Agd derived cyclic peptides, a
 CC composition containing a peptide and a carrier, and a method for the
 CC production of the cyclic peptides. The peptide inhibits the agr response,
 CC which is normally associated with the release of virulence factors of
 CC Staphylococcus aureus. An Agd peptide is produced by *S. aureus* that
 CC activates the agr response in strains of a single group, but interferes
 CC with this response in strains of different groups. The peptides and
 CC composition containing them can be used to treat infections by *S. aureus*.

[illegible]

```

Db          1 GVNASSSLF 9
          |||:|
RESULT 8
ID AAM51003 standard; Peptide; 9 AA.
XX
AC AAM51003;
XX
DE 08-MAY-2002 (first entry)
XX
DE AgrD2 linear free acid peptide.
XX
KW AgrD2; agr response; inhibitor; antibiotic; antibacterial;
XX infection; therapy.
XX
OS Staphylococcus aureus.
XX Synthetic.
XX
FN US6337385-B1.
XX
PD 08-JAN-2002.
XX
PF 24-JUN-1999; 99US-0339511.
XX
PR 24-JUN-1998; 98US-090402P.
XX
PA (UYRQ ) UNIV ROCKEFELLER.
XX (UYNY ) UNIV NEW YORK STATE.
XX
PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX
DR WPI; 2002-170774/22.
XX
PT Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
XX interference and for treating Staphylococcus aureus infection in a
XX subject -
XX
PS Example 1; Column 9; 18bp; English.
XX
CC The present sequence is that of a novel synthetic AgrD2 linear
XX free acid peptide. The peptide is derived from the cyclic AgrD2
XX peptide of Staphylococcus aureus group II. AgrD2 is a secreted
XX agr-encoded peptide, where the agr locus controls the synthesis of
XX virulence factor and other extracellular proteins responsible for
XX pathogenicity in S. aureus. The biological activity of the
XX synthetic peptide was assayed using cultured S. aureus strains
XX containing a beta-lactamase reporter gene fused to the agr
XX promoter. This allowed activation or inhibition of the agr
XX response to be monitored spectrophotometrically. Unlike an AgrD2
XX thiolactone cyclic peptide (see AAM51001), the present peptide
XX was unable to either activate or inhibit the agr response, even
XX when added to cultured cells at 1M concentrations. The invention
XX provides claimed cyclic peptides (see AAM50899-906 and AAM50999) for
XX and methods for preparing them. The cyclic peptides are useful for
XX bacterial interference, especially for the treatment of S. aureus
XX infection.
SQ Sequence 9 AA;

Query Match          89.7%; Score 35; DB 23; Length 9;
Best Local Similarity 77.8%; Pred. No. 9.3e+05;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXASLRF 9
   |||:|
Db 1 GVNASSSLF 9

RESULT 9
ID AAM51004 standard; Peptide; 9 AA.
          |||:|
XX
AC AAM51004;
XX
DE 08-MAY-2002 (first entry)
XX
DE AgrD2 lactone cyclic peptide.
XX
KW AgrD2; agr response; inhibitor; antibiotic; antibacterial;
XX infection; therapy; lactone; cyclic.
XX
OS Staphylococcus aureus.
XX Synthetic.
XX
FN US6337385-B1.
XX
PD 08-JAN-2002.
XX
PF 24-JUN-1999; 99US-0339511.
XX
PR 24-JUN-1998; 98US-090402P.
XX
PA (UYRQ ) UNIV ROCKEFELLER.
XX (UYNY ) UNIV NEW YORK STATE.
XX
PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX
DR WPI; 2002-170774/22.
XX
PT Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
XX interference and for treating Staphylococcus aureus infection in a
XX subject -
XX
PS Example 1; Column 9; 18bp; English.
XX
CC The present sequence is that of a novel synthetic AgrD2
XX lactone cyclic peptide in which residue 5 of the peptide is
XX linked to residue 9 via a lactone bond. The peptide is derived
XX from an AgrD2 peptide of Staphylococcus aureus group II. AgrD2 is
XX a secreted agr-encoded peptide, where the agr locus controls the
XX synthesis of virulence factor and other extracellular proteins
XX responsible for pathogenicity in S. aureus. The biological
XX activity of the synthetic peptide was assayed using cultured S.
XX aureus strains containing a beta-lactamase reporter gene fused to
XX the agr promoter. This allowed activation or inhibition of the
XX agr response to be monitored spectrophotometrically. The lactone
XX AgrD2 peptide inhibited the agr response of group I S. aureus
XX strains without activating the agr response in group I, II or III
XX strains. The invention provides claimed cyclic peptides (see
XX AAM50899-906 and AAM50999) and methods for preparing them,
XX especially peptides where the cyclic bond is a lactam or lactone
XX bond. The cyclic peptides are useful for bacterial interference,
XX especially for the treatment of S. aureus infection.
SQ Sequence 9 AA;

Query Match          89.7%; Score 35; DB 23; Length 9;
Best Local Similarity 77.8%; Pred. No. 9.3e+05;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXASLRF 9
   |||:|
Db 1 GVNASSSLF 9

RESULT 10

```

AB07160
 ID ABB07160 standard; peptide; 9 AA.
 XX
 AC ABB07160;
 XX
 DT 13-MAR-2002 (first entry)
 XX
 DE Peptide-mediated quorum sensing inhibitor peptide cyclo-XII.
 XX
 KM Autoinducer-2; A1-2; antibiotic; antibacterial; dermatological;
 KM vulnerability; pheromone; agr system; accessory gene regulator; cyclic.
 XX
 OS Synthetic.
 XX
 PN WO200185664-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 10-MAY-2001; 2001WO-US15221.
 XX
 PR 10-MAY-2000; 2000US-203000P.
 PR 07-DEC-2000; 2000US-254398P.
 XX
 PA (UVR-) UNIV PRINCETON.
 PA (QUOR-) QUORX PHARM INC.
 PA (UYE-) UNIV TECHNOLOGIES INT INC.
 PI Bassler BL, Dammel CS, Schauder S, Shokat K, Stein J, Surette MG;
 XX
 DR WPI; 2002-075235/10.
 XX
 PT Use of autoinducer-2 agonists or antagonists for regulating activity of
 PT autoinducer-2 receptor, regulating bacterial growth and pathogenesis.
 XX
 PS also antibiotic compositions
 XX
 PS Disclosure; Page 33; 134pp; English.
 XX
 CC The invention relates to the use of autoinducer-2 (A1-2) agonists or
 CC antagonists for regulating activity of autoinducer-2 receptor, regulating
 CC bacterial growth and pathogenesis. Synergistic antibiotic compositions
 CC comprising inhibitors of the quorum-sensing pathway of a microorganism
 CC are also provided. Methods using such A1-2 analogues are useful for
 CC treating pathogen-associated disease states. The compounds and antibiotic
 CC compositions can be used to inhibit bacterial cell growth and/or biofilm
 CC formation on a medical device, particularly for promoting growth of skin
 CC graft replacements used in the treatment of burns and ulcers. They may
 CC also be used to aid wound repair, and to inhibit bacterial cell growth
 CC and biofilm formation in or on products or devices used for personal
 CC hygiene. The present sequence represents a inhibitor of peptide-mediated
 CC quorum sensing.
 CC
 SQ Sequence 9 AA;
 XX
 Query Match 89.7%; Score 35; DB 23; Length 9;
 Best Local Similarity 77.8%; Pred. No. 9.3e+05;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GVNAXASLP 9
 Db 1 GVNASSSLF 9
 XX
 RESULT 11
 ID AAY67851 standard; peptide; 9 AA.
 XX
 AC AAY67851;
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE S. aureus peptide #1 used for bacterial interference.
 DE
 XX Staphylococcus aureus infection; cyclic peptide; AgrD; agr response;
 KM

virulence factor; treatment.
 XX
 OS Staphylococcus aureus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 5
 FT /label= Unknown
 XX
 PN WO9967286-A2.
 XX
 PD 29-DEC-1999.
 XX
 PF 24-JUN-1999; 99WO-US14562.
 XX
 PR 24-JUN-1998; 98US-0103438.
 XX
 PA (UVR-) UNIV ROCKEFELLER.
 PA (UTNY) UNIV NEW YORK STATE.
 XX
 PI Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
 XX
 DR WPI; 2000-147202/13.
 XX
 PT New cyclic peptides for treating infections with Staphylococcus aureus
 PT
 PS Claim 9; Page 26; 37pp; English.
 XX
 CC This sequence represents a cyclic peptide derived from the Staphylococcus
 CC aureus AgrD peptide. The invention relates to AgrD derived peptides,
 CC a composition containing a peptide and a carrier, and a method for the
 CC production of the cyclic peptides. The peptide inhibits the agr response,
 CC which is normally associated with the release of virulence factors of
 CC Staphylococcus aureus. An AgrD peptide is produced by S. aureus that
 CC activates the agr response in strains of a single group, but interferes
 CC with this response in strains of different groups. The peptides and
 CC composition containing them can be used to treat infections by S. aureus.
 CC
 SQ Sequence 9 AA;
 XX
 Query Match 87.2%; Score 34; DB 21; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVNAXASLP 9
 Db 1 GVNAXSSLP 9
 XX
 RESULT 12
 ID AAY67861 standard; peptide; 9 AA.
 XX
 AC AAY67861;
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE Staphylococcus aureus AgrDII derived peptide sequence.
 DE
 XX Staphylococcus aureus infection; AgrD; agr response; treatment;
 KM virulence factor.
 KM
 OS Staphylococcus aureus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 5
 FT /label= Unknown
 XX
 PN WO9967286-A2.
 XX
 PD 29-DEC-1999.
 XX
 PF 24-JUN-1999; 99WO-US14562.
 XX

XX 24-JUN-1998; 98US-0103438.
 PR (UTRO) UNIV ROCKEFELLER.
 XX (UTNY) UNIV NEW YORK STATE.
 PA
 XX Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
 PI WPI; 2000-147202/13.
 DR
 XX New cyclic peptides for treating infections with Staphylococcus aureus
 PT
 PS Examples; Page 22; 37pp; English.
 CC This sequence represents the Staphylococcus aureus AgrD derived
 CC peptide. The invention relates to AgrD derived cyclic peptides, a
 CC composition containing a peptide and a carrier, and a method for the
 CC production of the cyclic peptides. The peptide inhibits the agr response,
 CC which is normally associated with the release of virulence factors of
 CC Staphylococcus aureus. An AgrD peptide is produced by S. aureus that
 CC activates the agr response in strains of a single group, but interferes
 CC with this response in strains of different groups. The peptides and
 CC composition containing them can be used to treat infections by S. aureus.
 CC
 SQ Sequence 9 AA;
 Query Match 87.2%; Score 34; DB 23; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVNAXASLRF 9
 DB 1 GVNAXSSLRF 9
 RESULT 13
 ABP53540
 ID ABP53540 standard; peptide; 9 AA.
 XX
 AC ABP53540;
 XX
 DT 13-DEC-2002 (first entry)
 XX
 DE Cyclic peptide SEQ ID NO:1.
 XX
 KM Cyclic peptide; Staphylococcus aureus; infection; antibacterial;
 KW agr response inhibitor.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 5 /note= "any amino acid"
 XX
 FT US200207453-A1.
 XX
 PN 20-JUN-2002.
 XX
 PD 27-DEC-2001; 2001US-0032950.
 PF
 XX 24-JUN-1998; 98US-090402P.
 PR 24-JUN-1999; 99US-0339511.
 XX
 XX (Muir/) Muir T W.
 PA (Mayv/) Mayville P.
 PA (Novi/) Novick R P.
 PA (Beav/) Beavis R.
 PA (JiG/) Ji G.
 XX
 XX Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
 PI WPI; 2002-681366/73.

XX New cyclic peptides, useful for treating Staphylococcus aureus
 PT infections
 FT
 XX Claim 9; Page 10; 18pp; English.
 PS
 CC ABP53540 to ABP53547 represent cyclic peptides (I) from the present
 CC invention. The present invention also describes a method for treating
 CC Staphylococcus aureus infection comprising the administration of a
 CC composition comprising (I). (I) has antibacterial activity, and can be
 CC used as an agr gene response inhibitor. The peptides are useful for
 CC treating S. aureus infections.
 CC
 SQ Sequence 9 AA;
 Query Match 87.2%; Score 34; DB 23; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVNAXASLRF 9
 DB 1 GVNAXSSLRF 9
 RESULT 14
 AAM50899
 ID AAM50899 standard; Peptide; 9 AA.
 XX
 AC AAM50899;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE AgrD-autoinducing cyclic peptide, inhibitor of agr response.
 XX
 KM Staphylococcus aureus; AgrD; agr response; inhibitor; antibiotic;
 KW antibacterial; infection; therapy; cyclic.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 5 /note= "any amino acid"
 FT Misc-difference 5 /note= "note linked to residue 9 to form cyclic
 FT peptide"
 FT Misc-difference 9 /note= "note linked to residue 5 to form cyclic
 FT peptide"
 XX
 PN US6337385-B1.
 XX
 PD 08-JAN-2002.
 XX
 PF 24-JUN-1999; 99US-0339511.
 PR 24-JUN-1998; 98US-090402P.
 XX
 PA (UTRO) UNIV ROCKEFELLER.
 PA (UTNY) UNIV NEW YORK STATE.
 XX
 PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
 DR WPI; 2002-170774/22.
 XX
 XX Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
 PT interference and for treating Staphylococcus aureus infection in a
 PT subject
 FT
 XX Claim 7; Column 19; 18pp; English.
 PS
 CC The present sequence is that of a novel synthetic cyclic peptide
 CC of the invention that is capable of inhibiting the agr response of
 CC Staphylococcus aureus. It is an AgrD-autoinducing peptide, where

CC AgRD is a secreted agr-encoded peptide and where the agr locus
 CC controls the synthesis of virulence factor and other extracellular
 CC proteins responsible for pathogenicity in *S. aureus*. Preferred
 CC peptides may have the sequence NH₂-X(n)-Z-X(y)-COOH, with a
 CC cyclic bond between the Z residue and COOH other than a thioester
 CC bond, where X is an amino acid, an amino acid analogue, a
 CC peptidomimetic or non-amide isostere, Z is a synthetic or a
 CC biosynthetic amino acid, n is 0-10 and y is 1-10. The cyclic bond
 CC is especially a lactam or lactone bond. The thiololactone
 CC structure within native AgrD peptides is required for activation
 CC of the agr response. Elimination of the thiol ester component of
 CC the cyclic ring structure can destroy agr response activating
 CC activity while preserving and enhancing inhibitory activity. A
 CC claimed method of preparing a cyclic peptide involves: assembling
 CC a linear peptide chain on to a solid phase resin support;
 CC deprotecting the resulting protected assembled peptide; treating the
 CC deprotected peptide with neutral buffer for a time sufficient to
 CC form the cyclic peptide and cleave the peptide from the support;
 CC and recovering the cyclic peptide. The peptide is useful for
 CC bacterial interference, especially for the treatment of *S. aureus*
 CC infection.

CC Sequence 9 AA;

Query Match 87.2%; Score 34; DB 23; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVNAXASLRF 9
 1 GVNAXASLRF 9

RESULT 15
 AAMS0908
 ID AAMS0908 standard; Peptide: 9 AA.

AC AAMS0908;

DT 08-MAY-2002 (first entry)

DE Protected peptide used in cyclic peptide production.

KM *Staphylococcus aureus*; AgrD; agr response; inhibitor; antibiotic;

KM antibacterial; infection; therapy; cyclic.

OS Synthetic.

Key Location/Qualifiers

Modified-site 1 /note= "Z-Gly"

Modified-site 5 /label= Dpr(Boc)

Modified-site 6 /note= "Ser(Bz1)"

Modified-site 7 /note= "Ser(Bz1)"

Misc-difference 5 /note= "note linked to residue 9 to form cyclic peptide"

Misc-difference 9 /note= "note linked to residue 5 to form cyclic peptide"

US6337385-B1.

06-JAN-2002.

24-JUN-1999; 99US-0339511.

24-JUN-1998; 98US-090402P.

(UVRQ) UNIV ROCKEFELLER.

PA (UTNY) UNIV NEW YORK STATE.

PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;

WPI; 2002-170774/22.

Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
 interference and for treating *Staphylococcus aureus* infection in a
 subject

Disclosure; Column 14; 18pp; English.

The present sequence is that of a protected peptide used in an
 example of the preparation of novel synthetic cyclic peptides of
 the invention (see AAMS0899-906). The peptide corresponds to the
Staphylococcus aureus AgrDII sequence with a Cys5 to diaminopropionic
 acid (Dpr) mutation (lactam). It was synthesized on a Wang-resin
 using an Fmoc N-alpha protection strategy. Following chain assembly,
 the peptide was cleaved from the support and the Dpr-5 residue
 deprotected by treatment with a trifluoroacetic acid:anisole:water
 mixture (90:5:5) for 4 hr. The partially protected peptide-alpha
 carboxylates were then dissolved in DMF and treated with PyBOP.
 Cyclization was complete after 2 hr. The remaining protecting groups
 were removed by treatment with HF and the peptide purified by HPLC.
 The cyclic peptide is capable of inhibiting the agr response of
Staphylococcus aureus. The thiololactone structure within native
 AgrD peptides is required for activation of this response.
 Replacement of the thiol ester component of the cyclic ring
 structure with a lactam (as in the present case) or a lactone can
 destroy agr response activating activity while preserving and
 enhancing inhibitory activity. The cyclic peptides are useful for
 bacterial interference, especially for the treatment of *S. aureus*
 infection.

Sequence 9 AA;

Query Match 87.2%; Score 34; DB 23; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVNAXASLRF 9
 1 GVNAXASLRF 9

Search completed: January 29, 2004, 14:52:12
 Job time : 35.0735 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:46:49 ; Search time 11.1176 Seconds
(without alignments)
77.851 Million cell updates/sec

Title: US-10-032-950-5

Perfect score: 39

Sequence: 1 GYNAXSLF 9

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database:

PIR 76: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	94.9	333	1	C69812 ferrichrome ABC tr
2	33	84.6	47	2	Agd protein limpo
3	32	82.1	382	2	lysosomal membrane
4	32	82.1	405	2	P28/LAMP-1 precurs
5	32	82.1	407	2	120K lysosomal mem
6	31	79.5	349	2	S74439 iron(III) diclitr
7	31	79.5	468	2	T24523 hypothetical prote
8	30	76.9	188	2	B38116 hypothetical prote
9	30	76.9	188	2	transposase alr850
10	30	76.9	234	2	transposase alr850
11	30	76.9	247	2	ferric exochelin u
12	30	76.9	247	2	transposase alr700
13	30	76.9	247	2	transposase alr730
14	30	76.9	247	2	transposase alr711
15	30	76.9	247	2	transposase alr737
16	30	76.9	247	2	transposase alr710
17	30	76.9	247	2	transposase alr717
18	30	76.9	259	2	cytochrome-c oxida
19	30	76.9	346	2	iron(III) diclitr
20	30	76.9	361	2	hypothetical prote
21	30	76.9	458	2	probable glutamine
22	30	76.9	458	2	transposase alr732
23	30	76.9	497	2	transposase alr732
24	30	76.9	497	2	transposase alr732
25	30	76.9	497	2	transposase alr732
26	30	76.9	572	2	transposase alr726
27	30	76.9	573	2	urease (EC 3.5.1.5
28	30	76.9	649	2	urease (EC 3.5.1.5
29	30	76.9	1232	2	iron (III) diclitr
30	30	76.9	1232	2	multidrug resistanc
31	29	74.4	109	2	hypothetical prote

ALIGNMENTS

30	29	74.4	148	2	D97308	probable membrane
31	29	74.4	488	2	A53572	prostaglandin E2 r
32	29	74.4	500	1	EPF	zip protein precu
33	29	74.4	570	2	T33320	hypothetical prote
34	29	74.4	605	2	H71562	probable flagellar
35	29	74.4	655	2	G99189	protein Y33F6.3 l
36	29	74.4	703	2	T41065	RNA binding protei
37	29	74.4	895	2	T45786	receptor-protein k
38	29	74.4	1238	2	T40120	C2H2 type zinc fin
39	28	71.8	186	2	D83354	hypothetical prote
40	28	71.8	255	2	AD1159	flagellar biosynth
41	28	71.8	255	2	AD1518	flagellar biosynth
42	28	71.8	297	2	B81194	polymine transpor
43	28	71.8	333	2	D59805	iron(III) diclitr
44	28	71.8	369	2	A44942	fructose-bisphosph
45	28	71.8	369	2	B43715	protein M homolog

RESULT 1

C69812 ferrichrome ABC transporter (permease) homolog yfmd - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C/Accession: C69812

R/Kunze, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber

C.; Brion, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;

A.; Enlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,

Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Gal

leob, J.; Harwood, C.R.; Hensaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M

Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino

A.; Luch, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau

Y, M.; Ogasawara, K.; Ogiwara, A.; Onda, B.; Parx, S.H.; Parro, V.; Pohl, T.M.; Portet

Rieger, M.; Rivolta, C.; Roche, B.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scall

A.; Schaefer, S.; Schaefer, R.; Scott, F.; Sekiguchi, J.; Sekowska, A.; Se

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpetra, P.; Toigoni, A.; Toasato, V.; Uchiya

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamano, K.; Yasumoto, K.; Yara, K.; Yoshida

A.; Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Zumbstein, A.

A./Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A./Reference number: A69580; MUID:95044033; PMID:9384377

A./Accession: C69812

A./Status: preliminary; nucleic acid sequence not shown; translation not shown

A./Molecule type: DNA

A./Residues: 1-333 <KIN>

A./Cross-references: GB:Z99108; GB:AL009126; NID:G2633055; PIDN:CAB12580.1; PID:G26330

A./Experimental source: strain 168

A./Gene: yfmd

C/Genetics:

C/Superfamily: ferrichrome ABC transporter

Query Match 94.9%; Score 37; DB 1; Length 333;

Best Local Similarity 88.9%; Pred. 0.52;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYNAXSLF 9

Db 99 GYNAXSLF 107

RESULT 2

Agd protein (imported) - Staphylococcus aureus (strain N315)

C/Species: Staphylococcus aureus

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C/Accession: C89995

R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O

ma, A.; Mizutani, U.; Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K

C.; Shibata, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratake, K.

Lancet 357, 1225-1240, 2001

A./Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146
 A:Accession: C89995
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-47 <KDJ>
 A:Cross-references: GB:BA000018; FID:G13701831; PIDN:BA843124.1; GSPDB:GN00149
 A:Experimental source: strain N315
 C:Genetics:
 A:Gene: agtD

Query Match 84.6%; Score 33; DB 2; Length 47;
 Best Local Similarity 77.8%; Pred. No. 0.61;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXASLF 9
 |||||
 DB 24 GVNACSSSLF 32

RESULT 3

A28067
 A:lysosomal membrane glycoprotein LAMP-1 - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 20-Aug-1999
 C:Accession: A28067
 R:Chen, J.W.; Cha, Y.; Yukael, K.U.; Gracy, R.W.; August, J.T.
 J. Biol. Chem. 263, 8754-8758, 1988
 A:Title: Isolation and sequencing of a cDNA clone encoding lysosomal membrane glycoprotein
 A:Reference number: A28067; MUID:88243732; PMID:3379044
 A:Accession: A28067
 A:Molecule type: mRNA
 A:Residues: 1-382 <CHB>
 A:Cross-references: GB:J03881; NID:G198706; PIDN:AAA9411.1; PID:G293692
 A:Note: the authors translated the codon ATT for residue 1 as Leu and CCG for residue 2
 C:Superfamily: lysosome-associated membrane protein
 C:Keywords: glycoprotein; membrane protein

Query Match 82.1%; Score 32; DB 2; Length 382;
 Best Local Similarity 66.7%; Pred. No. 9.8;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXASLF 9
 |||||
 DB 256 GVNASSSLF 264

RESULT 4

A60534
 A:PB/LAMP-1 precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 03-May-1996
 C:Accession: A60534
 R:Hefferman, M.; Yousefi, S.; Dennis, J.W.
 Cancer Res. 49, 6077-6084, 1989
 A:Title: Molecular characterization of PB/LAMP-1, a major protein target of a metastasizing cell line
 A:Reference number: A60534; MUID:90002989; PMID:2676155
 A:Accession: A60534
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-405 <HEF>
 C:Superfamily: lysosome-associated membrane protein

Query Match 82.1%; Score 32; DB 2; Length 405;
 Best Local Similarity 66.7%; Pred. No. 10;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXASLF 9
 |||||
 DB 279 GVNASSSLF 267

RESULT 5

A30200

120K lysosomal membrane glycoprotein precursor - rat

N:Alternate names: statoglycoprotein
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 18-Oct-1989 #sequence_revision 30-Sep-1991 #text_change 20-Aug-1999
 C:Accession: A30200; S03331
 R:Howe, C.L.; Granger, B.L.; Hull, M.; Green, S.A.; Gabel, C.A.; Helentius, A.; Mellina
 Proc. Natl. Acad. Sci. U.S.A. 85, 7577-7581, 1988
 A:Title: Derived protein sequence, oligosaccharides, and membrane insertion of the 12
 glycoproteins.

A:Reference number: A30200; MUID:89017240; PMID:3174652
 A:Accession: A30200
 A:Molecule type: mRNA
 A:Residues: 1-407 <HIN>

A:Cross-references: EMBL:J03672
 A:Note: the authors translated the codon GGG for residue 15 as Val
 R:Himeno, M.; Noguchi, Y.; Sasaki, H.; Tanaka, Y.; Furuno, K.; Kono, A.; Sasaki, Y.;
 FEBS Lett. 244, 351-356, 1989
 A:Title: Isolation and sequencing of a cDNA clone encoding 107 kDa statoglycoprotein
 A:Reference number: S03331; MUID:89153580; PMID:2920835
 A:Accession: S03331

A:Molecule type: mRNA

A:Residues: 22-407 <HIN>
 A:Cross-references: EMBL:X14765; NID:956577; PIDN:CAA32873.1; PID:956578
 A:Note: part of this sequence, including the amino end of the mature protein, was con
 C:Superfamily: lysosome-associated membrane protein
 C:Keywords: glycoprotein; membrane protein

F:22-407/Product: lysosomal membrane glycoprotein, 107K #status experimental <MAT>

Query Match 82.1%; Score 32; DB 2; Length 407;
 Best Local Similarity 66.7%; Pred. No. 10;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXASLF 9
 |||||
 DB 281 GVNATSSLF 289

RESULT 6

S74439
 A:iron(III) dicitrate transport system permease protein fecd - Synecocystis sp. (strat
 N:Alternate names: protein slr1317
 C:Species: Synecocystis sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
 C:Accession: S74439
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
 O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocys
 S:
 A:Reference number: S74439; MUID:97061201; PMID:8905231
 A:Accession: S74439
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-349 <KAN>
 A:Cross-references: EMBL:D90899; GB:BA001339; NID:G1651650; PIDN:BA16591.1; PID:G165
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Genetics:
 A:Gene: fecd
 C:Superfamily: vitamin B12 transport protein btuc
 C:Keywords: iron transport

Query Match 79.5%; Score 31; DB 2; Length 349;
 Best Local Similarity 87.5%; Pred. No. 15;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXASLF 8
 |||||
 DB 113 GVNAGSLF 120

RESULT 7

T24523

hypothetical protein T05E11.5 - *Caenorhabditis elegans*
 C/Species: *Caenorhabditis elegans*
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T24523
 R/Kershaw, J.
 submitted to the EMBL Data Library, January 1996
 A/Reference number: Z19903
 A/Accession: T24523
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-468 <NT>
 A/Cross-references: EMBL:Z68751; PIDN:CAA92975.1; GSPDB:GN00022; CESP:T05E11.5
 A/Experimental source: clone T05E11
 A/Genetics:
 A/Gene: CESP:T05E11.5
 A/Map position: 4
 A/Introns: 23/3; 293/3

Query Match 79.5%; Score 31; DB 2; Length 468;
 Best Local Similarity 87.3%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAKXSLF 8
 |||||
 Db 181 GVNAFASL 188

RESULT 8
 B38116
 hypothetical protein 2 - *Anabaena* sp. insertion sequence IS892
 C/Species: *Anabaena* sp.
 C/Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 24-Nov-1999
 C/Accession: B38116
 R/Cai, Y.
 J. Bacteriol. 173, 5771-5777, 1991
 A/Title: Characterization of insertion sequence IS892 and related elements from the cyanobacterium *Anabaena*
 A/Reference number: A38116; MUID:91358369; PMID:1653218
 A/Accession: B38116
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-188 <CAI>
 A/Cross-references: GB:W64297
 A/Experimental source: PCC 7120
 C/Genetics:
 A/Mobile element: insertion sequence IS892
 C/Superfamily: *Anabaena* insertion sequence IS892 hypothetical protein 2

Query Match 76.9%; Score 30; DB 2; Length 188;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAKXSLF 9
 |||||
 Db 5 GLSATASLF 13

RESULT 9
 AC2563
 transposase a1r8502 [imported] - *Nostoc* sp. (strain PCC 7120) plasmid pCC7120delta
 C/Species: *Nostoc* sp. PCC 7120
 A/Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
 C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C/Accession: AC2563
 R/Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena*
 A/Reference number: AB1807; MUID:21595285; PMID:11759840
 A/Accession: AC2563
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-188 <KUR>
 A/Cross-references: GB:AP003604; PIDN:BAW7421.1; PID:G17134865; GSPDB:GN00183

A/Experimental source: strain PCC 7120
 C/Genetics:
 A/Gene: a1r8502
 A/Genome: plasmid
 C/Superfamily: *Anabaena* insertion sequence IS892 hypothetical protein 2

Query Match 76.9%; Score 30; DB 2; Length 188;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAKXSLF 9
 |||||
 Db 5 GLSATASLF 13

RESULT 10
 S60885
 ferric exochelin uptake protein fxuC - *Mycobacterium smegmatis* (fragment)
 C/Species: *Mycobacterium smegmatis*
 C/Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 22-Oct-1999
 C/Accession: S60885
 R/Fiss, E.H.; Yu, S.; Jacobs Jr., W.R.
 Mol. Microbiol. 14, 557-569, 1994
 A/Title: Identification of genes involved in the sequestration of iron in mycobacteria
 A/Reference number: S60885; MUID:95191405; PMID:7885234
 A/Accession: S60885
 A/Molecule type: DNA
 A/Residues: 1-234 <FIS>
 A/Cross-references: EMBL:U10425; NID:9595400; PIDN:AAQ3258.1; PID:9595401
 C/Genetics:
 A/Gene: fxuC

Query Match 76.9%; Score 30; DB 2; Length 234;
 Best Local Similarity 55.6%; Pred. No. 18;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAKXSLF 9
 |||||
 Db 37 GINSKALF 45

RESULT 11
 AD2478
 transposase a117004 [imported] - *Nostoc* sp. (strain PCC 7120) plasmid pCC7120alpha
 C/Species: *Nostoc* sp. PCC 7120
 A/Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
 C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C/Accession: AD2478
 R/Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena*
 A/Reference number: AB1807; MUID:21595285; PMID:11759840
 A/Accession: AD2478
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-247 <KUR>
 A/Cross-references: GB:BA000020; PIDN:BAW78088.1; PID:G1713542; GSPDB:GN00180
 C/Genetics:
 A/Mobile element: insertion sequence IS892 hypothetical protein 2
 C/Superfamily: *Anabaena* insertion sequence IS892 hypothetical protein 2

Query Match 76.9%; Score 30; DB 2; Length 247;
 Best Local Similarity 66.7%; Pred. No. 19;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAKXSLF 9
 |||||
 Db 64 GLSATASLF 72

RESULT 12

AF2515

transposase all1302 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120alpha

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C:Accession: AF2515

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; PMID:21595285; PMID:11759840

A:Accession: AF2515

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-247 <KUR>

A:Cross-references: GB:BA000020; PIDN:BAB78386.1; PID:G17135840; GSPDB:GN00180

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all1302

A:Genome: plasmid

C:Superfamily: Anabaena insertion sequence IS892 hypothetical protein 2

Query Match

Best Local Similarity	66.7%	Score 30;	DB 2;	Length 247;
Matches	6;	Conservative	2;	Mismatches 1;
				Indels 0;
				Gaps 0;

QY 1 GVNAXASLF 9

DB 64 GLSATASLF 72

RESULT 13

AG2491

transposase all1111 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120alpha

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C:Accession: AG2491

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; PMID:21595285; PMID:11759840

A:Accession: AG2491

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-247 <KUR>

A:Cross-references: GB:BA000020; PIDN:BAB78195.1; PID:G17135649; GSPDB:GN00180

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all1111

A:Genome: plasmid

C:Superfamily: Anabaena insertion sequence IS892 hypothetical protein 2

Query Match

Best Local Similarity	76.9%	Score 30;	DB 2;	Length 247;
Matches	6;	Conservative	2;	Mismatches 1;
				Indels 0;
				Gaps 0;

QY 1 GVNAXASLF 9

DB 64 GLSATASLF 72

RESULT 14

AG2524

transposase all1735 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120alpha

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C:Accession: AG2524

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; PMID:21595285; PMID:11759840

A:Accession: AG2524

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-247 <KUR>

A:Cross-references: GB:BA000020; PIDN:BAB77133.1; PID:G17134574; GSPDB:GN00180

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all1735

A:Genome: plasmid

C:Superfamily: Anabaena insertion sequence IS892 hypothetical protein 2

Query Match

Best Local Similarity	76.9%	Score 30;	DB 2;	Length 247;
Matches	6;	Conservative	2;	Mismatches 1;
				Indels 0;
				Gaps 0;

QY 1 GVNAXASLF 9

DB 64 GLSATASLF 72

RESULT 15

AT2480

transposase all17105 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120alpha

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C:Accession: AT2480

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; PMID:21595285; PMID:11759840

A:Accession: AT2480

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-247 <KUR>

A:Cross-references: GB:BA000020; PIDN:BAB78189.1; PID:G17135643; GSPDB:GN00180

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all17105

A:Genome: plasmid

C:Superfamily: Anabaena insertion sequence IS892 hypothetical protein 2

Query Match

Best Local Similarity	76.9%	Score 30;	DB 2;	Length 247;
Matches	6;	Conservative	2;	Mismatches 1;
				Indels 0;
				Gaps 0;

QY 1 GVNAXASLF 9

DB 64 GLSATASLF 72

Search completed: January 29, 2004, 14:58:22

Job time : 11.1176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:45:44 ; Search time 5.82353 Seconds

(without alignments)
72.678 Million cell updates/sec

Title: us-10-032-950-5

Perfect score: 39

Sequence: 1 GVMAXSLF 9

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Length	ID	Description
1	82.1	406	1 LMP1_MOUSE	P11438 mus musculu
2	82.1	407	1 LMP1_RAT	P11562 ratu
3	79.5	468	1 YJ95_CAEEL	P49049 caenorhabdi
4	76.9	571	1 URE1_YEREN	P31494 yersinia en
5	76.9	571	1 URE1_YERPE	O94f13 yersinia pe
6	76.9	571	1 URE1_YERPE	P52313 yersinia pe
7	74.4	468	1 PE24_HUMAN	P31408 homo sapien
8	74.4	468	1 UZP1_DROME	P13279 drosophila
9	71.8	369	1 ALP_PLAFA	P14223 plasmodium
10	71.8	369	1 MX_STRPY	P15946 streptococc
11	71.8	386	1 SELA_HELIPY	O26043 helicobacte
12	71.8	389	1 KSS5_ECOLI	P42218 escherichia
13	71.8	390	1 RBM9_HUMAN	O43251 homo sapien
14	71.8	407	1 LMP1_RIGR	P48129 cricetus
15	71.8	631	1 YBBS_BACST	P38616 bacillus su
16	71.8	631	1 SC24_YEAST	P49482 saccharomyc
17	71.8	2363	1 SPCO_MOUSE	Q62261 mus musculu
18	71.8	2364	1 SPCO_MOUSE	Q01082 homo sapien
19	69.2	169	1 COMD_METJA	P58415 methanococc
20	69.2	197	1 VS11_ROTMA	P45416 human rocv
21	69.2	230	1 TRBP_SYNEL	O84085 synecococc
22	69.2	244	1 YSCD_YEREN	Q01251 yersinia en
23	69.2	244	1 YSCD_YERPE	Q00926 yersinia pe
24	69.2	256	1 TRYB_MANSE	P35045 manduca sex
25	69.2	256	1 TRYB_MANSE	P35046 manduca sex
26	69.2	256	1 TRYB_CHOPU	P35042 chorisicneu
27	69.2	334	1 FEED_ECOLI	P21876 escherichia
28	69.2	340	1 CBRB_ERMCH	O47085 ewinia chr
29	69.2	439	1 RT30_HUMAN	Q94992 homo sapien
30	69.2	442	1 RT30_MOUSE	Q94090 mus musculu
31	69.2	446	1 MOXY_PARDE	P25905 paracoccus
32	69.2	465	1 NORA_CHLMD	Q991u3 chlamydia m
33	69.2	465	1 NORA_CHLTR	O84639 chlamydia t

ALIGNMENTS

RESULT 1
ID LMP1_MOUSE STANDARD; PRT; 406 AA.
AC P11438; Q62020; (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 15, Last sequence update)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Lysosome-associated membrane glycoprotein 1 precursor (LAMP-1) (LGP-A)
DE (LGP-120) (CD107A) (P25).
GN LAMP1 OR LAMP-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=50307338; PubMed=2142158;
RA Granger B.L., Green S.A., Gabel C.A., Howe C.L., Mellman I.,
RA Helenius A.;
RT "Characterization and cloning of 19p10, a lysosomal membrane
RT glycoprotein from mouse and rat cells."
RT J. Biol. Chem. 265:12036-12043 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=50307338; PubMed=2142158;
RA Granger B.L., Green S.A., Gabel C.A., Howe C.L., Mellman I.,
RA Helenius A.;
RT "Characterization and cloning of 19p10, a lysosomal membrane
RT glycoprotein from mouse and rat cells."
RT J. Biol. Chem. 265:12036-12043 (1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=50307338; PubMed=2142158;
RA Granger B.L., Green S.A., Gabel C.A., Howe C.L., Mellman I.,
RA Helenius A.;
RT "Characterization and cloning of 19p10, a lysosomal membrane
RT glycoprotein from mouse and rat cells."
RT J. Biol. Chem. 265:12036-12043 (1990).
RN [4]
RP DISULFIDE BONDS.
RX MEDLINE=90237040; PubMed=2332434;
RA Arterburn L.M., Earles B.J., August J.T.;
RT "The disulfide structure of mouse lysosome-associated membrane
RT protein 1."
RT J. Biol. Chem. 265:7419-7423 (1990).
CC - FUNCTION: PRESENTS CARBOHYDRATE LIGANDS TO SELECTINS. ALSO
CC IMPLICATED IN TUMOR CELL METASTASIS.
CC - SUBCELLULAR LOCATION: Type I membrane protein. Lysosomal.
CC THIS PROTEIN SHUTTLES BETWEEN LYSOSOMES, ENDOSOMES, AND THE PLASMA
CC MEMBRANE.
CC - PTM: O- AND N-GLYCOSYLATED, SOME OF THE N-GLYCANS ATTACHED TO
CC LAMP-1 ARE POLYLACTOSAMINOGLYCANS (BY SIMILARITY).
CC - SIMILARITY: BELONGS TO THE LAMP FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

RESULT 2		
LMP1_RAT		
ID	LMP1_RAT	STANDARD; PRT: 407 AA.
AC	P14362; P97620;	
DT	01-JAN-1990 (Rel. 13, Created)	
DT	01-JAN-1990 (Rel. 13, Last sequence update)	
DT	01-NOV-1997 (Rel. 35, Last annotation update)	
DE	Lysosome-associated membrane glycoprotein 1 precursor (LAMP-1) (20 kDa lysosomal membrane glycoprotein) (LGP-120) (CD107A).	
GN	LAMP1 OR LAMP-1.	
OC	Rattus norvegicus (Rat)	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI TaxID=10116;	

Accession	Source	Length	Score	Ident	Accession	Source	Length	Score	Ident					
FN	SEQUENCE FROM N.A.				FN	SEQUENCE FROM N.A.								
RP	MEDLINE=89017240; PubMed=3174652;				RP	MEDLINE=89017240; PubMed=3174652;								
RX	Howe C.T., Granger B.L., Hull M., Green S.A., Gabel C.A., Helenius A.,				RX	Howe C.T., Granger B.L., Hull M., Green S.A., Gabel C.A., Helenius A.,								
RA	Mellman I.;				RA	Mellman I.;								
RA	"Derived protein sequence, oligosaccharides, and membrane insertion				RA	"Derived protein sequence, oligosaccharides, and membrane insertion								
RT	of the 120-kDa lysosomal membrane glycoprotein (lgp120);				RT	of the 120-kDa lysosomal membrane glycoprotein (lgp120);								
RT	identification of a highly conserved family of lysosomal membrane				RT	identification of a highly conserved family of lysosomal membrane								
RT	glycoproteins.";				RT	glycoproteins.";								
RL	Proc. Natl. Acad. Sci. U.S.A. 85:7577-7581(1988).				RL	Proc. Natl. Acad. Sci. U.S.A. 85:7577-7581(1988).								
RN	[2]				RN	[2]								
RP	SEQUENCE OF 283-357 FROM N.A.				RP	SEQUENCE OF 283-357 FROM N.A.								
RA	MEDLINE=89153580; PubMed=2920635;				RA	MEDLINE=89153580; PubMed=2920635;								
RA	Hirano M., Noguchi Y., Sasaki H., Tanaka Y., Futuno K., Kono A.,				RA	Hirano M., Noguchi Y., Sasaki H., Tanaka Y., Futuno K., Kono A.,								
RT	Sakaki Y., Kato K.;				RT	Sakaki Y., Kato K.;								
RT	"Isolation and sequencing of a cDNA clone encoding 107 kDa				RT	"Isolation and sequencing of a cDNA clone encoding 107 kDa								
RL	salivary glycoprotein in rat liver lysosomal membranes.";				RL	salivary glycoprotein in rat liver lysosomal membranes.";								
RL	FEBS Lett. 244:351-356(1989).				RL	FEBS Lett. 244:351-356(1989).								
CC	[3]				CC	[3]								
RP	SEQUENCE OF 283-357 FROM N.A.				RP	SEQUENCE OF 283-357 FROM N.A.								
RC	STRAIN=Wistar Kyoto; TISSUE=Arctic smooth muscle;				RC	STRAIN=Wistar Kyoto; TISSUE=Arctic smooth muscle;								
RL	Adams L.A., Mehry I., Schwartz S.M.;				RL	Adams L.A., Mehry I., Schwartz S.M.;								
CC	Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.				CC	Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.								
CC	-1- FUNCTION: PRESENTS CARBOHYDRATE LIGANDS TO SELECTINS. ALSO				CC	-1- FUNCTION: PRESENTS CARBOHYDRATE LIGANDS TO SELECTINS. ALSO								
CC	IMPLICATED IN TUMOR CELL METASTASIS.				CC	IMPLICATED IN TUMOR CELL METASTASIS.								
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein. Lysosomal.				CC	-1- SUBCELLULAR LOCATION: Type I membrane protein. Lysosomal.								
CC	THIS PROTEIN SHUTTLES BETWEEN LYSOSOMES, ENDOSOMES, AND THE PLASMA				CC	THIS PROTEIN SHUTTLES BETWEEN LYSOSOMES, ENDOSOMES, AND THE PLASMA								
CC	MEMBRANE.				CC	MEMBRANE.								
CC	-1- PM: O- AND N-GLYCOSYLATED. SOME OF THE N-GLYCANS ATTACHED TO				CC	-1- PM: O- AND N-GLYCOSYLATED. SOME OF THE N-GLYCANS ATTACHED TO								
CC	LAMP-1 ARE POLYLACTOSAMINOGLYCANS (BY SIMILARITY).				CC	LAMP-1 ARE POLYLACTOSAMINOGLYCANS (BY SIMILARITY).								
CC	-1- SIMILARITY: BELONGS TO THE LAMP FAMILY.				CC	-1- SIMILARITY: BELONGS TO THE LAMP FAMILY.								
CC	-----				CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				CC	between the Swiss Institute of Bioinformatics and the EMBL outstation								
CC	at the European Bioinformatics Institute. There are no restrictions on its				CC	at the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way				CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial				CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/</td><td></td><td></td><td></td></tr> <tr> <td>CC</td><td>or send an email to license@isb-sib.ch).				CC	or send an email to license@isb-sib.ch).			
CC	-----				CC	-----								
DR	EMBL; M34959; AAA41525.1; -				DR	EMBL; M34959; AAA41525.1; -								
DR	EMBL; X14765; CAA32873.1; -				DR	EMBL; X14765; CAA32873.1; -								
DR	EMBL; U75406; AAB19108.1; -				DR	EMBL; U75406; AAB19108.1; -								
DR	PIR; A30200; A30200.				DR	PIR; A30200; A30200.								
DR	InterPro: IPR002000; LAMP.				DR	InterPro: IPR002000; LAMP.								
DR	Pfam; PF01293; LAMP; 1				DR	Pfam; PF01293; LAMP; 1								
DR	PRINTS; PR00336; LYSASOCTMP.				DR	PRINTS; PR00336; LYSASOCTMP.								
DR	PROSITE; PS00310; LAMP_1; 2.				DR	PROSITE; PS00310; LAMP_1; 2.								
DR	PROSITE; PS00311; LAMP_2; 1.				DR	PROSITE; PS00311; LAMP_2; 1.								
KW	Transmembrane; Glycoprotein; Lysosome; Signal.			</										


```

FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 339 339 N -> T (IN REF. 3).
FT CONFLICT 356 357 SD -> VT (IN REF. 3).
SQ SEQUENCE 407 AA; 43969 MW; 2594790749A7C88 CRC64;

Query Match 82.1%; Score 32; DB 1; Length 407;
Best Local Similarity 66.7%; Pred. No. 5.2;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXASLF 9
DB 281 GVNATISLF 289

RESULT 3
Y055-CABEL STANDARD; PRT; 468 AA.
ID Y055-CABEL
AC P49049;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 52.8 kDa protein T05B11.5 in chromosome IV.
GN T05B11.5
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Kershaw G.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO YEAST YKL100C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z68751; CAA92975.1; -
DR PIR; T24523; T24523.
DR WormPep; T05B11.5; CE06364.
DR InterPro; IPR006639; PSN.
DR Pfam; PF04258; DUF435; 1.
DR SMART; SM00730; PSN; 1.
DR SMART; SM00730; PSN; 1.
KM Hypothetical protein.
SQ SEQUENCE 468 AA; 52793 MW; 00DFACB679A5F8 CRC64;

Query Match 79.5%; Score 31; DB 1; Length 468;
Best Local Similarity 67.5%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXASLF 8
DB 181 GVNAPASLF 188

RESULT 4
URE1_YEREN STANDARD; PRT; 571 AA.
ID URE1_YEREN
AC P31494;
DT 01-JUL-1993 (Rel. 26, Created)

```

```

DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN URE1_YEREN
OS URE1_YEREN
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A2635 / Serotype O:8;
RC MEDLINE=94320783; PubMed=8045421;
RA de Koning-Ward T.F., Ward A.C., Robins-Browne R.M.;
RT "Characterisation of the urease-encoding gene complex of Yersinia
RT enterocolitica."
RL Gene 145:25-32(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=6471/76 / Serotype O:3;
RC MEDLINE=93273501; PubMed=8500886;
RA Skurnik M., Batsford S., Metz A.K.H., Schilz E., Toivanen P.;
RT "The putative arthritogenic cationic 19-kilodalton antigen of
RT Yersinia enterocolitica is a urease beta-subunit."
RL Infect. Immun. 61:2498-2504(1993).
CC -1- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -1- COFACTOR: Binds 2 nickel ions per subunit (By similarity).
CC -1- SUBUNIT: (ALPHA, BETA, GAMMA)3 (BY SIMILARITY).
CC -1- PTM: Lys-221 is carbamylated. The carbamoyl group provides the
CC 13ands for the two nickel ions (By similarity).
CC -1- SIMILARITY: BELONGS TO THE UREASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; L24101; AAA50996.1; -
DR EMBL; Z18665; CAA79216.1; -
DR PIR; S36028; S36028.
DR HSSP; P18314; LFWF.
DR MEROPS; M38.UNM; -
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005847; Urease.
DR InterPro; IPR005848; UreaseA.
DR Pfam; PF01979; Amidohydro_1; 1.
DR Pfam; PF00449; Urease1; 1.
DR PROSITE; PS00145; Urease_2; 1.
DR PROSITE; PS01120; UREASE_1; 1.
KM Hydrolyase; Metal-binding; Nickel.
FT INIT MET 0
FT METAL 138 138 NICKEL 2 (BY SIMILARITY).
FT METAL 140 140 NICKEL 2 (BY SIMILARITY).
FT METAL 221 221 NICKEL 1 AND 2 (BY SIMILARITY).
FT METAL 250 250 NICKEL 1 (BY SIMILARITY).
FT METAL 276 276 NICKEL 2 (BY SIMILARITY).
FT METAL 364 364 NICKEL 2 (BY SIMILARITY).
FT ACT SITE 324 324 BY SIMILARITY.
FT CONFLICT 29 30 IE -> MQ (IN REF. 2).
FT CONFLICT 66 66 V -> GY (IN REF. 2).
FT CONFLICT 68 82 DLVITNVTVDARFG -> EFSHNQRHYCSPFR
FT CONFLICT 96 96 A -> V (IN REF. 2).
FT CONFLICT 114 114 M -> L (IN REF. 2).
FT CONFLICT 137 137 S -> T (IN REF. 2).
FT CONFLICT 405 405 L -> V (IN REF. 2).
FT CONFLICT 440 440 D -> V (IN REF. 2).
SQ SEQUENCE 571 AA; 60952 MW; 99C7944554FB58EA CRC64;

Query Match 76.9%; Score 30; DB 1; Length 571;
Best Local Similarity 66.7%; Pred. No. 22;

```

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSLP 9
Db 309 GVNQAEFL 317

RESULT 5
ID UREL_YERPE STANDARD; PRT; 571 AA.
AC Q92FR9;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolyase).
GN UREAC Q92567 OR Y1239.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6/59M;
RA Sebanae F., Devalckeneere A., Simonet M.;
RT "Characterization of the urease locus from Yersinia pestis";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RA MEDLINE=21470413; PubMed=1156360;
RA Parkhill J., Wren B.W., Thomson N.R., Titchall R.M., Holden M.T.G.,
RA Prentice M.B., Sebanae M., James K.D., Churcher C., Mangai K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Farraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltham T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,
RA Leach S., Moul S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague";
RL Nature 411:523-527(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RA MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liles P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM";
RL J. Bacteriol. 184:4601-4611(2002).
CC -1- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -1- COFACTOR: Binds 2 nickel ions per subunit (By similarity).
CC -1- SUBUNIT: (ALPHA, BETA, GAMMA) (3) (By similarity).
CC -1- PTM: Lys-221 is carbamylated. The carbamoyl group provides the
CC ligands for the two nickel ions (By similarity).
CC -1- SIMILARITY: BELONGS TO THE UREASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF095636; AAC78634.1; -
CC EMBL: AJ414153; CAC92906.1; -
CC EMBL: AF013727; AAM64814.1; ALT_INIT.
CC PIR: AC0325; AC0325.
CC HSSP: P18314; IFW.
CC InterPro: IPR006680; Amidohydro_1.
CC InterPro: IPR005847; Urease.
CC InterPro: IPR005848; Ureasea.

DR Pfam: PF01979; Amidohydro_1; 1.
DR Pfam: PF00449; urease; 1.
DR PROSITE: PS00145; UREASE_2; 1.
DR PROSITE: PS01120; UREASE_1; 1.
KM Hydrolyase; Metal-binding; Nickel: Complete proteome.
FT INIT_MET 0
FT METAL 138 138 BY SIMILARITY
FT METAL 140 140 NICKEL 2 (BY SIMILARITY).
FT METAL 221 221 NICKEL 1 AND 2 (BY SIMILARITY).
FT METAL 250 250 NICKEL 1 (BY SIMILARITY).
FT METAL 276 276 NICKEL 1 (BY SIMILARITY).
FT METAL 364 364 NICKEL 2 (BY SIMILARITY).
FT ACT_SITE 324 324 BY SIMILARITY
SQ SEQUENCE 571 AA; 60888 MW; D35D233B75EBB26 CRC64;
Query Match 76.9%; Score 30; DB 1; Length 571;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSLP 9
Db 309 GVNQAEFL 317

RESULT 6
ID UREL_YERPS STANDARD; PRT; 571 AA.
AC P52313;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolyase).
GN UREAC
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP 27751;
RA MEDLINE=97270511; PubMed=9125594;
RA Rioc B., Berche P., Simonet M.;
RT "Urease is not involved in the virulence of Yersinia
RT pseudotuberculosis in mice";
RL Infect. Immun. 65:1985-1990(1997).
CC -1- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -1- COFACTOR: Binds 2 nickel ions per subunit (By similarity).
CC -1- SUBUNIT: (ALPHA, BETA, GAMMA) (3) (By similarity).
CC -1- PTM: Lys-221 is carbamylated. The carbamoyl group provides the
CC ligands for the two nickel ions (By similarity).
CC -1- SIMILARITY: BELONGS TO THE UREASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U40842; AAA87854.1; -
CC HSSP: P18314; IFW.
CC MEROPS: M38 UNM; -
CC InterPro: IPR005847; Urease.
CC InterPro: IPR006680; Amidohydro_1.
CC InterPro: IPR005848; Ureasea.
CC Pfam: PF01979; Amidohydro_1; 1.
CC Pfam: PF00449; urease; 1.
CC PROSITE: PS00145; UREASE_2; 1.
CC PROSITE: PS01120; UREASE_1; 1.
KM Hydrolyase; Metal-binding; Nickel.
FT INIT_MET 0
FT METAL 138 138 BY SIMILARITY
FT METAL 138 138 NICKEL 2 (BY SIMILARITY).

```

FT METAL 140 140 NICKEL 2 (BY SIMILARITY).
FT METAL 221 221 NICKEL 1 AND 2 (BY SIMILARITY).
FT METAL 250 250 NICKEL 1 (BY SIMILARITY).
FT METAL 276 276 NICKEL 1 (BY SIMILARITY).
FT METAL 364 364 NICKEL 2 (BY SIMILARITY).
FT ACT SITE 324 324 BY SIMILARITY.
SQ SEQUENCE 571 AA; 60874 MW; 88B6056FA16F89CC CRC64;

Query Match 76.9%; Score 30; DB 1; Length 571;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXASLF 9
Db 309 GVNQSASLF 317

RESULT 7
PE24 HUMAN STANDARD; PRT; 488 AA.
ID PE24 HUMAN
AC P35408;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Prostaglandin E2 receptor, EP4 subtype (Prostanoid EP4 receptor) (PGE
DE receptor, EP4 subtype).
GN PTGER4 OR PTGER2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94216291; PubMed=8163486;
RA Bastien L., Sawyer N., Gygoczyk R., Metters K.M., Adam M.;
RT "Cloning, functional expression, and characterization of the human
RT prostaglandin E2 receptor EP2 subtype.";
RL J. Biol. Chem. 269:11873-11877(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94071953; PubMed=8250933;
RA An S., Yang J., Xia M., Goetzl E.J.;
RT "Cloning and expression of the EP2 subtype of human receptors for
RT prostaglandin E2.";
RL Biochem. Biophys. Res. Commun. 197:263-270(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96299755; PubMed=8661119;
RA Ford S.M., Marks B., Stolz M., Buffler E., Fraser N.J., Lee M.G.;
RT "The structure of the prostaglandin EP4 receptor gene and related
RT pseudogenes.";
RL Genomics 35:182-188(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=97015877; PubMed=8662514;
RA Mori K., Tanaka I., Kotani M., Miyake F., Sando T., Muro S.,
RA Sasaki Y., Nakagawa O., Ogawa Y., Usui T., Ozaki S., Ichikawa A.,
RA Natsumiya S., Nakao K.;
RT "Gene expression of the human prostaglandin E receptor EP4 subtype:
RT differential regulation in monocyctoid and lymphoid lineage cells by
RT phorbol ester.";
RL J. Mol. Med. 74:333-336(1996).
RN [5]
RP FUNCTION: RECEPTOR FOR PROSTAGLANDIN E2 (PGE2). THE ACTIVITY OF
RN [6]
RP ADENYLYL CYCLASE. HAS A RELAXING EFFECT ON SMOOTH MUSCLE. MAY
RN [7]
RP PLAY AN IMPORTANT ROLE IN REGULATING RENAL HEMODYNAMICS,
RN [8]
RP INTERSTITIAL EPITHELIAL TRANSPORT, ADRENAL ALDOSTERONE SECRETION,
RN [9]
RP AND UTERINE FUNCTION.
RN [10]
RP SUBCELLULAR LOCATION: Integral membrane protein.
RN [11]
RP TISSUE SPECIFICITY: HIGH IN INTESTINE AND IN PERIPHERAL BLOOD

```

```

CC MONONUCLEAR CELLS; LOW IN LUNG, KIDNEY, THYMUS, UTERUS,
CC VASCULATURE AND BRAIN. NOT FOUND IN LIVER, HEART, RETINA OR
CC SKELETAL MUSCLE.
CC - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC - CAUTION: Was originally designated as the EP2 subtype.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, L28175; AAA36434.1; -
CC EMBL, L25124; AAA36438.1; -
CC EMBL, X97873; CAA66463.1; -
CC EMBL, X97874; CAA66463.1; JOINED.
CC EMBL, D28472; BAA05834.1; -
CC PIR, A53572; A53572.
CC Genew; HGNC:9596; PTGER4.
CC MIM; 601586; -
CC GO; GO:0016021; C: integral to membrane; TAS.
CC GO; GO:0004957; F: prostaglandin E receptor activity; TAS.
CC GO; GO:0007188; P: G-protein signaling, coupled to camp nucleo. .; TAS.
CC GO; GO:0006955; P: immune response; TAS.
CC InterPro; IPR00276; GPCR_Rhodopsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCR_RHODOPSIN.
CC PROSITE; PS00337; G-PROTEIN RECP_F1_1; 1.
CC PROSITE; PS00262; G-PROTEIN RECP_F1_2; 1.
CC KX G-protein coupled receptor; Transmembrane; Glycoprotein.
CC FT DOMAIN 1 19
CC FT TRANSMEM 20 43
CC FT DOMAIN 44 55
CC FT TRANSMEM 56 79
CC FT DOMAIN 80 96
CC FT TRANSMEM 97 115
CC FT DOMAIN 116 135
CC FT TRANSMEM 136 160
CC FT DOMAIN 161 184
CC FT TRANSMEM 185 211
CC FT DOMAIN 212 267
CC FT TRANSMEM 268 295
CC FT DOMAIN 296 312
CC FT TRANSMEM 313 332
CC FT DOMAIN 333 488
CC FT CARBOHYD 7 7
CC FT DISULFID 92 170
CC FT DOMAIN 230 234
CC FT CONFLICT 464 466
CC FT SEQUENCE 488 AA; 53119 MW; D028478CD72C85EB CRC64;

Query Match 74.4%; Score 29; DB 1; Length 488;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXASLF 8
Db 5 GVNQSASLF 12

RESULT 8
UZIP DROME STANDARD; PRT; 488 AA.
ID UZIP DROME
AC P10379; Q960B4; Q9M0W7;
DT 01-MAR-1989 (Rel. 10, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Unipped protein precursor (zipper protein).
GN UZIP OR ZIP OR CG3533.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=68296414; PubMed=3402433;
 RA Zhao D.-B., Code S., Jaehnis F., Haller J., Jaechle H.;
 RT "zipper encodes a putative integral membrane protein required for
 RL normal axon patterning during Drosophila neurogenesis.";
 RL EMBO J. 7:1115-1119(1988).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaratunga C., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V., Chapple M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Bencs P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Jengam C.,
 RA Jaiswal B., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 RA Kimmel B.B., Kinsler P., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasco P., Lei Y., Levitsky A.A., Li Z., Li Z., Liang Y., Lin X.,
 RA Liu X., Martell B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K.A., Nuskern D., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 [3]
 RP REVISIONS.
 RC STRAIN=Berkeley;
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Mungall J.S., Milburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Celniker S.E., de Grey A.D.N.T., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RL systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley; TISSUE=Embryo;
 RX MEDLINE=22426066; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champagne M.,
 RA George R.A., Guarin H., Krommiller B., Pacle J.M., Park S., Wan K.H.,
 RA Rubin G.M., Celniker S.E.;

RT "A Drosophila full-length cDNA resource.";
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 CC -1- FUNCTION: Required for normal axon patterning during neurogenesis.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; X07450; CAA30332.1; -;
 DR EMBL; AE003465; AAF47312.2; -;
 DR EMBL; AY052139; AAK93563.1; -;
 DR PIR; S00483; EPFF.
 DR FlyBase; FBgn0004055; uzpD.
 KM Differentiation; Neurogenesis; Transmembrane; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 488
 FT DOMAIN 22 465
 FT TRANSMEM 466 486
 FT DOMAIN 487 488
 FT DOMAIN 380 400
 FT CARBOHYD 35 35
 FT CARBOHYD 232 232
 FT CARBOHYD 317 317
 FT CARBOHYD 374 374
 FT CARBOHYD 448 448
 FT CONFLICT 488 488
 FT SEQUENCE 488 AA; 54308 MW; 5B4E93D9012A2AB7 CRC64;
 Query Match 74.4%; Score 29; DB 1; Length 488;
 Best Local Similarity 66.7%; Freq. No. 33;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GVNXASLF 9
 DB 463 GVNXASLF 471
 ID ALF PLFA STANDARD; PRT; 369 AA.
 AC P14223;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Fructose-bisphosphate aldolase (EC 4.1.2.13) (41 kDa antigen).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OC NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90136715; PubMed=2693962;
 RA Knapp B., Hundt B., Kuepper H.A.;
 RT "A new blood stage antigen of Plasmodium falciparum transported to
 RL the erythrocyte surface.";
 RL Mol. Biochem. Parasitol. 37:47-56(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90271928; PubMed=2190085;
 RA Knapp B., Hundt B., Kuepper H.A.;
 RT "Plasmodium falciparum aldolase: gene structure and localization.";
 RL Mol. Biochem. Parasitol. 40:1-12(1990).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
 RX MEDLINE=98190013; PubMed=9521758;
 RA Kim H., Cetta U., Dobeil H., Jakob P., Hol W.G.U.;
 RT "Crystal structure of fructose-1,6-bisphosphate aldolase from the
 RL human malaria parasite Plasmodium falciparum.";
 RL Biochemistry 37:4388-4396(1998).

```

FT HELIX 284 295
FT TURN 296 296
FT STRAND 303 306
FT HELIX 310 320
FT TURN 321 322
FT HELIX 324 326
FT HELIX 327 345
FT TURN 346 347
SQ SEQUENCE 369 AA; 401.05 MW; 2AE9CDEDAF5C96A4 CRC64;

Query Match
Best Local Similarity 66.7%; Score 28; DB 1; Length 369;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYNAXASLF 9
DB 355 GENAGASLY 363

RESULT 10
ID MX_STRPY STANDARD; PRT; 369 AA.
AC P16946;
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Virulence factor-related M protein precursor.
GN ENNX.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
CX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=serotype M49;
RX MEDLINE=90078078; PubMed=2687231;
RA Haanes S.J., Cleary P.P.;
RT "Identification of a divergent M protein gene and an M protein-related
   gene family in Streptococcus pyogenes serotype 49.",
   J. Bacteriol. 171:6337-6408(1989).
RL -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
   an amide bond (Potential).
CC -1- SIMILARITY: BELONGS TO THE M PROTEIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
-----
CC EMBL, M23688; AAA26919.1; -.
CC EMBL, M31790; AAA26875.1; -.
CC PIR, B43715; B43715.
CC InterPro: IPR005877; Gpos_Y5IRK.
CC InterPro: IPR001899; Gram_pos_anchor.
CC InterPro: IPR006192; LPXTG.
CC InterPro: IPR003345; M_repeat.
CC Pfam: PF00746; Gram_pos_anchor; 1.
CC Pfam: PF02370; M; 7.
CC Pfam: PF04650; Y5IRK_signal; 1.
CC PRINTS: PR00015; GPOSANCHOR.
CC TIGRPFAMS: TIGR01167; LPXTG_anchor; 1.
CC TIGRPFAMS: TIGR01168; Y5IRK_signal; 1.
CC PROSITE: PSS0847; GRAM_POS_ANCHORING; 1.
CC Virulence, Phagocytosis; Cell wall, Peptidoglycan-anchor; Repeat;
CC Antigen; Colled coil; Signal.
CC SIGNAL 1 41
FT CHAIN 42 339 VIRULENCE FACTOR-RELATED M PROTEIN.
FT PROPEP 340 369 REMOVED BY SORTASE (POTENTIAL).
FT DOMAIN 129 200 2 X REPEATS, TYPE A.
FT REPEAT 129 159 A-1.

```

	CC	REPEAT	171	200	A-2	X REPEATS,	TYPE B.
	FT	DOMAIN	132	241	B-1.		
	FT	REPEAT	132	148	B-1.		
	FT	REPEAT	173	190	B-2.		
	FT	REPEAT	224	241	B-3.		
	FT	DOMAIN	150	215	C-1.		
	FT	REPEAT	150	160	C-1.		
	FT	REPEAT	192	201	C-2.		
	FT	REPEAT	206	215	C-3.		
	FT	DOMAIN	302	335	GLY/PRO-RICH (CELL WALL-SPANNING).		
	FT	SITE	336	340	LPMX SORTING SIGNAL (POTENTIAL).		
	FT	MOD RES	339	339	AMIDE-LINKED TO CELL WALL (POTENTIAL).		
	SEQUENCE	369 AA;	40639 MW;	4AD6BDA72F637839 CRCC4;			
	Query Match		71.8%;	Score 28;	DB 1;	Length 369;	
	Best Local Similarity	75.0%;	Pred. No. 43;				
	Matches	6;	Conservative	0;	Mismatches	2;	Indels 0; Gaps 0;
Oy		1 GVVATSL 8					
Dc		43 GVAVTSL 50					
RESULT 11							
ID	SELA_HELPY	STANDARD;	PRT;	386 AA.			
AC	026043;						
DT	16-OCT-2001 (Rel. 40, Created)						
DT	16-OCT-2001 (Rel. 40, Last sequence update)						
DT	28-FEB-2003 (Rel. 41, Last annotation update)						
DE	L-seryl-tRNA(Sec) selenium transferase (EC 2.9.1.1) (Cysteiny]-tRNA(Sec) selenium transferase) [Selenocysteine synthase]						
DE	(Selenocysteinyl-tRNA(Sec) synthase).						
DE	SELA OR HP1513.						
OS	Helicobacter pylori (Campylobacter pylori).						
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;						
OC	Helicobacteraceae; Helicobacter.						
OX	NCLTaxID=210;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=26695 / ATCC 700392;						
RX	MEDLINE=97394467; PubMed=9252185;						
RA	Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G., Frieschmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A., Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A., Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S., Loftus B., Richardson D., Dodson R., Kiratka H.G., Glodek A., McEnteny K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K., Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M., Cotton M.D., Wordeman J.M., Fujii C., Bowman C., Watney L., Wallin B., Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M., Venter J.C.;						
RA	Venter J.C.;						
RT	"the complete genome sequence of the gastric pathogen Helicobacter pylori."						
RU	Nature 388:539-547(1997)						
CC	- FUNCTION: Converts seryl-tRNA(Sec UCA) to selenocysteinyl-tRNA(Sec UCA) during selenoprotein biosynthesis (By similarity).						
CC	- CATALYTIC ACTIVITY: L-seryl-tRNA(Sec) + selenophosphate = L-selenocysteinyl-tRNA(Sec) + H ₂ O + phosphate.						
CC	- COPFACTOR: Pyridoxal phosphate (By similarity).						
CC	- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).						
CC	- SIMILARITY: BELONGS TO THE SELA FAMILY.						
CC	-----						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcement send an email to license@ebi.ac.uk).						
CC	EMBL, AE000649; AAC00554.1; -.						
DR	PIR, A64709; A64709.						
DR	TIGR, HP1513; -.						

```

DR HAMAP; MF_00423; -; 1.
DR InterPro; IPR004534; Sela.
DR Pfam; PF03841; Sela; 1.
KW Transferrase; Pyridoxal phosphate; Selenium; Complete proteome.
FT BINDING 221 221 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 386 AA; 43995 MW; 8ACDDE65073E859 CRC64;

Query Match
Best Local Similarity 71.8%; Score 28; DB 1; Length 386;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

CY 2 VANAXASL 9
      |||||
Db 21 INBASV 28

RESULT 12
KSS5_ECOLI STANDARD; PRT; 389 AA.
ID KSS5_ECOLI
AC P42218;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Capsule polysaccharide export protein kpsS.
GN kpsS.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
CX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K5;
RX MEDLINE=83368530; PubMed=8397187;
RA Pazzani C., Rosenow C., Boulnois G.J., Bromner D., Jann K.,
RA Roberts I.S.;
RT "Molecular analysis of region 1 of the Escherichia coli K5 antigen
RT gene cluster: a region encoding proteins involved in cell surface
RT expression of capsular polysaccharide.";
RL J. Bacteriol. 175:5978-5983(1993).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC CC EMBL; X74567; CAA52659.1; -
CC CC PIR; S36653; S36653.
CC CC Pfam; PF05159; Capsule_synth; 1.
CC CC Polysaccharide transport; Transport.
CC CC SEQUENCE 389 AA; 46381 MW; 73058122C28027DE CRC64;

Query Match
Best Local Similarity 75.0%; Score 28; DB 1; Length 389;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

CY 1 GVNAXASL 8
      |||||
Db 127 GVNAXSSL 134

RESULT 13
RM99_HUMAN STANDARD; PRT; 390 AA.
ID RM99_HUMAN
AC O43251; Q8TDO0; Q8WYB1; Q96DZ6; Q96NI7; Q9UGM4; Q9UH33;
DT 16-OCT-2001 (Rel. 40, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE RNA-binding protein 9 (RNA binding motif protein 9)
DE (Hearshonucleic acid binding protein 2) (Repressor of tamoxifen
DE transcriptional activity).

```


FT	CARBOHYD	102	102	N-LINKED (GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	116	116	N-LINKED (GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	125	125	N-LINKED (GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	145	145	N-LINKED (GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	160	160	N-LINKED (GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	178	178	N-LINKED (GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	215	215	N-LINKED (GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	220	220	N-LINKED (GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	233	233	N-LINKED (GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	241	241	N-LINKED (GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	253	253	N-LINKED (GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	283	283	N-LINKED (GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	297	297	N-LINKED (GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	304	304	N-LINKED (GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	312	312	N-LINKED (GLCNAC . .)	(POTENTIAL) .
SQ	SEQUENCE	407 AA;	43766 MW;	651002040F6BEB3D CRC64;	

Query Match:	71.8%;	Score 28;	DB 1;	Length 407;
Best Local Similarity	55.6%;	Pred. No. 48;		
Matches	5; Conservative	2; Mismatches	2; Indels	0; Gaps

Dn	1 GNMAXSLF 9			
	: : :			
	281 GNGSSSLF 289			


```

RESULT 15
YBFS_BACSU
ID YBFS_BACSU STANDARD; PRT; 631 AA.
AC P39616;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative PTS system IIBC component ybfS (EC 2.7.1.69).
GN ybfS.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1423;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Azevedo V., Barreiro M.G., Besieres P., Bolotin A., Borcher S.,
RA Borriss R., Bouvier L., Brans A., Braun M., Bignonell V., Carter N.M.,
RA Brouillet S., Bruschi C.V., Caldwell B., Cepano N.J., Daniel R.A.,
RA Choi S.K., Codan J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denicot F., Devine K.M., Duesterhoft A., Ehrlich S.D., Emerson P.T.,
RA Eutican K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
RA Fritz S.Y., Glaser P., Goffeau A., Golightly E.J., Grand G.,
RA Giuseppe G., Guy B.J., Haga K., Hsieh U., Harwood C.R., Hernat A.,
RA Horst H., Holzapfel S., Hosono S., Hullio M.F., Itaya M., Jones L.,
RA Jorls B., Karmali D., Kasahara Y., Kleier-Blanchard M., Klein C.,
RA Kobayashi Y., Koeller P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidos A., Lardinots S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Maeda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Ondega B., Park S.H.,
RA Parro V., Poh T.W., Portetelle D., Perwollik S., Prescott A.M.,
RA Pressac E., Puyc P., Puntelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scottone F.,
RA Sekiuchi J., Sekowska A., Serot S.J., Serrot P., Shin B.S., Soldo B.,
RA Sorokin A., Tagcon E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Tempstra P., Tognoni A.,

```


RA Tosato V., Uchiyama S., Vanderbol M., Vannier F., Vaesarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.",
RL Nature 390:249-256(1997).
RN (3)
RP SEQUENCE OF 515-631 FROM N.A.
RC STRAIN=168 / GCM;
RX MEDLINE=95270606; PubMed=7751298;
RA Tojner B., Uebink-Kox T., Poolman B., Konings W.N.;
RT "Characterization of the proton/glutamate symport protein of Bacillus
RT subtilis and its functional expression in Escherichia coli.",
RL J. Bacteriol. 177:2863-2869(1995).
CC -1- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHENOLPYRUVATE-DEPENDENT
CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE
CC -TRANSPORT SYSTEM. THE IID DOMAINS CONTAIN THE SUGAR BINDING SITE
CC AND THE TRANSMEMBRANE CHANNEL. THE IIA DOMAIN CONTAINS THE PRIMARY
CC PHOSPHORYLATION SITE (THE DONOR IS PHOSPHO-HPR); IIA TRANSFERS ITS
CC PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO
CC THE SUGAR (BY SIMILARITY).
CC CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
CC histidine + sugar phosphate.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Contains 1 PTS EIIA domain.
CC -1- SIMILARITY: Contains 1 PTS EIIB domain.
CC -1- SIMILARITY: Contains 1 PTS EIIC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB006424; BAA33132.1; -;
DR EMBL: 299105; CAH2029.1; -;
DR EMBL: U15147; AAA82877.1; -;
DR PIR: D69750; D69750.
DR HSSP: P20166; IGPR.
DR Subtilist; BG11014; YDFS.
DR InterPro; IPR001127; PTS_EIIA.
DR InterPro; IPR001996; PTS_EIIB.
DR InterPro; IPR003352; PTS_EIIC.
DR Pfam; PF00358; PTS_EIIA_1; 1.
DR Pfam; PF00367; PTS_EIIB; 1.
DR Pfam; PF02378; PTS_EIIC; 1.
DR ProDom; PD002243; PTS_EIIA; 1.
DR ProDom; PD001476; PTS_EIIB; 1.
DR TIGRfam; TIGR00826; EIIA_gic; 1.
DR TIGRfam; TIGR00830; PTER; 1.
DR PROSITE; PS00371; PTS_EIIA_1; 1.
DR PROSITE; PS01035; PTS_EIIB_CYS; 1.
KW Hypothetical protein; Phosphotransferase system; Sugar transport;
KW Transferrase; Phosphorylation; Transmembrane; Complete proteome.
FT MOD_RES 203 203 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 419 419 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 567 567 PHOSPHORYLATION (BY SIMILARITY).
FT DOMAIN 1 477 EIIC DOMAIN.
FT DOMAIN ? 477 EIIB DOMAIN.
FT DOMAIN 478 631 EIIA DOMAIN.
FT TRANSSEM 12 32 POTENTIAL.
FT TRANSSEM 56 76 POTENTIAL.
FT TRANSSEM 77 97 POTENTIAL.
FT TRANSSEM 106 126 POTENTIAL.
FT TRANSSEM 149 169 POTENTIAL.
FT TRANSSEM 173 193 POTENTIAL.
FT TRANSSEM 196 216 POTENTIAL.
FT TRANSSEM 243 263 POTENTIAL.
FT TRANSSEM 274 294 POTENTIAL.
FT TRANSSEM 298 318 POTENTIAL.

FT TRANSSEM 350 370 POTENTIAL.
SQ SEQUENCE 631 AA; 68144 MW; 883889A72EBC85B9B CRC64;
Query Match 71.8%; Score 28; DB 1; Length 631;
Best Local Similarity 62.5%; Pred. No. 76;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 GYNAXASTL 8
DB 172 GINAXASTL 179
Search completed: January 29, 2004, 14:53:10
Job time: 5.82353 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:46:19 ; Search time 26.8676 Seconds
(without alignments)
86.441 Million cell updates/sec

Title: US-10-032-950-5

Perfect score: 39

Sequence: 1 GVNXASLUF 9

Scoring table: BLOSUM62

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP viirus:*
16: SP bacteriap:*
17: SP archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	94.9	333	16	Q34933 bacillus su
2	33	84.6	47	16	Q33586 staphylococ
3	33	84.6	463	16	Q8XHH9 clostridium
4	32	82.1	189	11	Q92279 mus musculu
5	32	82.1	259	8	Q8NH50 Q8NH50 drugia mala
6	32	82.1	271	16	Q8BHVX3 Q8BHVX3 streptococ
7	32	82.1	406	11	Q8VH34 mus musculu
8	32	82.1	407	11	Q9DC13 Q9DC13 mus musculu
9	32	82.1	490	6	Q95XZ0 Q95XZ0 pan troglod
10	32	82.1	1248	6	Q9M838 arabidopsi
11	31	79.5	278	16	Q98X81 rhizobium1
12	31	79.5	349	16	P72591 synecocyst
13	31	79.5	578	16	Q9PKY0 chlamydia m
14	30	76.9	63	2	Q9RCK4 streptococ
15	30	76.9	188	16	Q9YK27 anabaena sp
16	30	76.9	234	2	Q50375 mycobacteri

17	30	76.9	242	8	Q94U25 Q94U25 leishmania
18	30	76.9	247	16	Q8YKCS Q8YKCS anabaena sp
19	30	76.9	247	16	Q8XGAI Q8XGAI anabaena sp
20	30	76.9	255	13	Q98SE3 Q98SE3 carassius a
21	30	76.9	259	8	Q47576 Q47576 onchocerca
22	30	76.9	337	16	Q8XW78 Q8XW78 clostridium
23	30	76.9	347	16	Q8ETK0 Q8ETK0 oceanobacil
24	30	76.9	346	16	Q8YTX8 Q8YTX8 anabaena sp
25	30	76.9	361	16	Q8UK10 Q8UK10 agrobacteri
26	30	76.9	458	16	Q916J5 Q916J5 pseudomonas
27	30	76.9	458	16	Q91275 Q91275 pseudomonas
28	30	76.9	497	16	Q8XFF5 Q8XFF5 anabaena sp
29	30	76.9	649	16	Q9XE20 Q9XE20 bacillus ha
30	30	76.9	1203	10	Q8S2E4 Q8S2E4 cryza sativ
31	30	76.9	1232	10	Q23998 Q23998 hordeum vul
32	30	76.9	1234	10	Q8G081 Q8G081 cryza sativ
33	29	74.4	46	2	Q9F1U5 Q9F1U5 staphylococ
34	29	74.4	103	6	Q9GMM8 Q9GMM8 macaca fasc
35	29	74.4	109	17	Q9UXG2 Q9UXG2 sulfobus
36	29	74.4	147	2	P77893 P77893 moraxella
37	29	74.4	148	16	Q97DZ7 Q97DZ7 clostridium
38	29	74.4	196	16	Q9CLC1 Q9CLC1 pasteurella
39	29	74.4	209	16	Q8XSM4 Q8XSM4 ralteonla s
40	29	74.4	310	5	Q61913 Q61913 toxocara ca
41	29	74.4	355	16	Q986V0 Q986V0 rhizobium1
42	29	74.4	385	5	Q8SRV6 Q8SRV6 encephalit
43	29	74.4	411	2	Q8KND1 Q8KND1 pseudomonas
44	29	74.4	411	2	Q9XC62 Q9XC62 pseudomonas
45	29	74.4	448	5	Q95XY8 Q95XY8 caenorhabd

ALIGNMENTS

RESULT 1	034933	PRELIMINARY;	PRT;	333	AA.
ID	034933				
AC	034933;				
DT	01-JAN-1998 (TREMBLrel. 05, Created)				
DT	01-JAN-1998 (TREMBLrel. 05, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)				
DE	YFMD protein.				
GN	YFMD.				
OS	Bacillus subtilis.				
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
OX	NCBI_TaxID=1423;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=168;				
RX	MEDLINE=98044033; PubMed=9384377;				
RA	Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,				
RA	Azevedo V., Besterio M.G., Bessieres P., Bolojin A., Borchert S.,				
RA	Boriss R., Boursier L., Brans A., Braun M., Bridnell S.C., Bron S.,				
RA	Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,				
RA	Choi S.K., Codani U.J., Connerton I.F., Cummings N.J., Daniel R.A.,				
RA	Denzot F., Devine K.M., Dusterhoft A., Enlrich S.D., Emmeson P.T.,				
RA	Entian K.D., Brington J., Fabre C., Ferrari E., Foulger D.,				
RA	Fritz C., Fujita M., Fujita S., Fuma S., Gallizi A., Galleron N.,				
RA	Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,				
RA	Ghissep G., Guy B.J., Haga K., Hasech J., Harwood C.R., Henaut A.,				
RA	Hilbert H., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,				
RA	Joris B., Karamata D., Kashara Y., Kleier-Blanchard M., Klein C.,				
RA	Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,				
RA	Kurita K., Lapidus A., Lardinois S., Lauber U., Lazarevic V.,				
RA	Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,				
RA	Medina M., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,				
RA	Moore D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,				
RA	Parro V., Pohl T.M., Portetelle B., Porrolik S., Prescott A.M.,				
RA	Preascan E., Puig P., Purnelle B., Porrolik S., Prescott A.M.,				
RA	Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadele Y.,				
RA	Sato T., Saitan E., Schlecht S., Schroeter R., Scottone F.,				
RA	Sekituchi Y., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,				
RA	Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,				

```

RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Toraochi V., Uchiyama S., Vandenberg M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wegler E., Wegler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H., Zunnstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=168;
RC Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=AC327;
RX MEDLINE=97417488; PubMed=9272861;
RA Yamamoto H., Uchiyama S., Nigroho F.A., Sekiguchi J.;
RT "Cloning and sequencing of a 35.7 Kb in the 70 degree-73 degree region
RT of the Bacillus subtilis genome reveal genes for a new two-component
RT system, three spore germination proteins, an iron uptake system and a
RT general stress response protein."
RL Gene 194:191-199(1997).
DR EMBL; Z99108; CAB12580.1; -
DR EMBL; D86417; BAA22318.1; -
DR InterPro; IPR000522; FecCD.
DR Pfam; PF01032; FecCD; 1.
DR ProDom; PD001557; FecCD; 1.
KM Complete proteome.
SQ SEQUENCE 333 AA; 35098 MW; 9F4BE91BDBDBE6761 CRC64;

Query Match 94.5%; Score 37; DB 16; Length 333;
Best Local Similarity 88.5%; Pred. No. 3.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps

Cy 1 GVNAXASLF 9
Db 99 GVNAGASLF 107

RESULT 2
033586 PRELIMINARY; PRT; 47 AA.
AC 033586;
DT 01-JAN-1998 (TrEMBLrel_05, Created)
DT 01-JAN-1998 (TrEMBLrel_05, Last sequence update)
DT 01-MAY-2002 (TrEMBLrel_20, Last annotation update)
DE AGRD (AGRD protein).
GN AGRD OR SAV2037 OR SA1842.1 OR SAS066.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
CX NCBI_TaxId=158878, 158879, 1280;
RN [1]
RN SEQUENCE FROM N.A.
RP SPECIES=S.aureus; STRAIN=SAS02A;
RC MEDLINE=97342847; PubMed=9197262;
RX Ji G., Beavie R., Novick R.P.;
RT "Bacterial interference caused by autoinducing peptide variants."
RL Science 276:2027-2030(1997).
RN [2]
RN SEQUENCE FROM N.A.
RP SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RC MEDLINE=2131195; PubMed=1148146;
RX Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-O., Ito T.,
RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hoshiyama A.,
RA Mitani-Uji Y., Takehashi N., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hiraoka H., Kunara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yanaishita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus

```

```

RT   Lancet 357:1225-1240(2001).
DR   EMBL: AF001782; BAB63265.1. -.
DR   EMBL: AP003364; BAB56199.1. -.
DR   EMBL: AP003135; BAB43124.1. -.
KW   Complete proteome.
SQ   SEQUENCE 47 AA; 5149 MW; 22893764DF54BFA7 CRC64;

Query Match
Best Local Similarity 84.6%; Score 33; DB 16; Length 47;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GYNAXASLP 9
DB 24 GYNACSLP 32

RESULT 3
O8XHH9 PRELIMINARY; PRT; 463 AA.
ID O8XHH9
AC O8XHH9;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypotheetical protein CPE2505.
GN CPE2505.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
CX NCBI_TaxID=1502;
CY [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX PubMed=11792842;
RA Shimitu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater".
RL Proc. Natl. Acad. Sci. U.S.A. 99:1996-1001(2002).
DR EMBL: AP003194; BAB82211.1. -.
DR InterPro: IPR002528; MATE.
DR Pfam: PF01554; MATE; 2.
DR TIGRFAMs: TIGR00797; MATE; 1.
KW Hypotheetical protein; Complete proteome.
SQ SEQUENCE 463 AA; 51100 MW; 6EE3B385633DEB9 CRC64;

Query Match
Best Local Similarity 84.6%; Score 33; DB 16; Length 463;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0.

OY 1 GYNAXASLP 9
DB 65 GYNACSLP 73

RESULT 4
ID O922T9 PRELIMINARY; PRT; 189 AA.
AC O922T9
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to lysosomal membrane glycoprotein 1 (Fragmant).
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC006785; AA006785.1. -.
DR InterPro: IPR002000; Lamp.

```

DR Pfam; PF01299; Lamp; 1.
 DR PRINTS; PR00336; LYSASOCTDMP.
 DR PROSITE; PS00310; LAMP_1; 1.
 DR PROSITE; PS00311; LAMP_2; 1.
 FT NON TER
 SQ SEQUENCE 189 AA; 20456 MW; 7AF16D69BA4066FA CRC64;

Query Match
 Best Local Similarity 82.1%; Score 32; DB 11; Length 189;
 Best Local Similarity 66.7%; Pred. No. 27;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXASLF 9
 DB 63 GNMASSLF 71

RESULT 5
 Q8HN50 PRELIMINARY; PRT; 259 AA.
 AC Q8HN50; 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Cytochrome c oxidase subunit III.
 GN COX3.
 OS Brugia malayi.
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Brugia.
 NCBI_TaxID=6279;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TRS;
 RA Dab U., Mudge J., Blaxter M.L.; The mitochondrial genome of the nematode parasite Brugia malayi.;
 RT Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF538716; AN17809.1; -
 DR EMBL; AF538716; AN17809.1; -
 KW Mitochondrion.
 SQ SEQUENCE 259 AA; 30791 MW; 3ABEE6464980596 CRC64;

Query Match
 Best Local Similarity 82.1%; Score 32; DB 8; Length 259;
 Best Local Similarity 66.7%; Pred. No. 37;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXASLF 9
 DB 127 GINGMASLF 135

RESULT 6
 Q8DVX3 PRELIMINARY; PRT; 271 AA.
 AC Q8DVX3; 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Conserved hypothetical protein, putative membrane protein.
 GN SMU.337.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UA159 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186;
 RA Adic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
 RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
 DR EMBL; AE014881; AAN58096.1; -

KW Hypothetical protein, Complete proteome.
 SQ SEQUENCE 271 AA; 31326 MW; 39AAB88F1444564 CRC64;

Query Match
 Best Local Similarity 82.1%; Score 32; DB 16; Length 271;
 Best Local Similarity 77.8%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXASLF 9
 DB 120 GVNPFASLF 128

RESULT 7
 Q8VH34 PRELIMINARY; PRT; 406 AA.
 AC Q8VH34; 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE LAMP-1.
 GN Mus musculus (Mouse).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ong K., Han J.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY069968; AAU58070.1; -
 DR InterPro; IPR002000; Lamp.
 DR Pfam; PF01299; Lamp; 1.
 DR PRINTS; PR00336; LYSASOCTDMP.
 DR PROSITE; PS00310; LAMP_1; 2.
 DR PROSITE; PS00311; LAMP_2; 1.
 DR PROSITE; PS00311; LAMP_2; 1.
 SQ SEQUENCE 406 AA; 43879 MW; C1BD373548ADF85 CRC64;

Query Match
 Best Local Similarity 82.1%; Score 32; DB 11; Length 406;
 Best Local Similarity 66.7%; Pred. No. 60;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXASLF 9
 DB 280 GNMASSLF 288

RESULT 8
 Q9DC13 PRELIMINARY; PRT; 407 AA.
 AC Q9DC13; 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Lysosomal membrane glycoprotein 1.
 GN LAMP1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; Tissue=Lung;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa W., Nishi K., Kiyosawa H., Kondo S., Yamana T.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Kono H., Baldarelli P., Barsh G.,
 RA Blake J., Botfeill D., Boujunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gardinot M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Togo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilmink L.,
 RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection,"
 RL Nature 409:685-690(2001).
 DR EMBL: AK004637; BAB23428.1; -
 DR MGD: MGI:96745; Lamp1.
 DR InterPro: IPR002000; Lamp.
 DR Pfam: PF01299; Lamp; 1.
 DR PRINTS: PR00336; LYSASOCTDMP.
 DR PROSITE: PS00310; LAMP_1; 2.
 DR PROSITE: PS00311; LAMP_2; 1.
 SO SEQUENCE 407 AA; 43936 MW; 696D0C79F627DA84 CRC64;

Query Match 82.1%; Score 32; DB 11; Length 407;
 Best Local Similarity 86.7%; Pred. No. 61;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GVNAXASLF 9
 Db 281 GVNAXASLF 289

RESULT 9
 ID Q9SKZ0 PRELIMINARY; PRT; 490 AA.
 AC Q9SKZ0;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Prostaglandin E2 subtype EP4 receptor.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
 OX NCBI_Taxid=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smock S.L., Castleberry T.A., Lu B., Owen T.A.,
 RT "Pan troglodytes prostaglandin E2 subtype EP4 receptor,"
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY052641; AAL15039.1; -
 DR InterPro: IPR00276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECPT_F1_1; 1.
 DR PROSITE: PS0262; G_PROTEIN_RECPT_F1_2; 1.
 KW Receptor.
 SO SEQUENCE 490 AA; 53303 MW; 2074ACAE9CBF87D0 CRC64;

Query Match 82.1%; Score 32; DB 6; Length 490;
 Best Local Similarity 87.5%; Pred. No. 74;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GVNAXASL 8
 Db 5 GVNAXASL 12

RESULT 10
 ID Q9M838 PRELIMINARY; PRT; 1248 AA.
 AC Q9M838;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE T27C4.14 protein.
 GN T27C4.14.
 OS Arabidopsis thaliana (House-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SRRAIN-cv. Columbia;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
 RA Walit R., Renning C.M., Koo H., Fujii C.Y., Utechtack T.R.,
 RA Rasthead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.,
 RT "Arabidopsis thaliana chromosome III BAC T27C4 genomic sequence,"
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC022287; AAF63780.1; -
 DR InterPro: IPR004314; DUF239.
 DR Pfam: PF03080; DUF239; 1.
 SO SEQUENCE 1248 AA; 138025 MW; 429E298908820456 CRC64;

Query Match 82.1%; Score 32; DB 10; Length 1248;
 Best Local Similarity 66.7%; Pred. No. 2e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GVNAXASLF 9
 Db 611 GVNAXASLF 619

RESULT 11
 ID Q98K81 PRELIMINARY; PRT; 279 AA.
 AC Q98K81;
 DT 01-OCT-2001 (TRENBLrel. 18, Created)
 DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
 DE Oxidoreductase of short-chain.
 GN MR1595.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_Taxid=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF30309;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asanizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.,
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti,"
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AF002897; BAB48933.1; -
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 DR PRINTS: PR00080; SDRFAMILY.
 KW Complete proteome.
 SO SEQUENCE 279 AA; 29917 MW; D32F8AFACBB1DB1 CRC64;

Query Match 79.5%; Score 31; DB 16; Length 279;
 Best Local Similarity 87.5%; Pred. No. 69;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GVNAXASL 8
 Db 139 GVNAXASL 146

RESULT 12
 ID P72591 PRELIMINARY; PRT; 349 AA.
 AC P72591;
 DT 01-FEB-1997 (TRENBLrel. 02, Created)
 DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)

DE IRON(III) dicitrate transport system permease protein FECD.
 GN FECD OR SLR1317.
 OC Synecocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
 RN NCBI_TaxID=1148;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirose A., Sugita M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naro K., Okumura S.,
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions";
 RL DNA Res. 3:109-136(1996).
 DR EMBL; D90899; BAA16591.1; -
 DR InterPro; IPR000522; FeccD.
 DR Pfam; PF01032; FeccD; 1.
 DR ProDom; PD001557; FeccD; 1.
 KM Complete proteome.
 SQ SEQUENCE 349 AA; 36770 MW; C077B37A52A897E1 CRC64;
 QY Query Match 79.5%; Score 31; DB 16; Length 349;
 Best Local Similarity 87.5%; Pred. No. 88;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 113 GVNAXASL 8
 113 GVNAXASL 120

RESULT 13
 Q9PKY0 PRELIMINARY; PRT; 578 AA.
 ID Q9PKY0;
 AC Q9PKY0;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Protein export protein, FHIPEP family.
 GN TC0330.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 RN NCBI_TaxID=83560;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mopn / N199;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey B.K., Peterson J., Utterback T., Berry K., Baas S.,
 RA Linde K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AR39";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 DR EMBL; AE002301; AAF73547.1; -
 DR TIGR; TC0330; -
 DR InterPro; IPR01712; Bact_FHIPEP.
 DR Pfam; PF00771; FHIPEP; 1.
 DR PRINTS; PR00948; TYPE1MAPROT.
 KM Complete proteome.
 SQ SEQUENCE 578 AA; 64153 MW; BAC9C4852D8B7F5C CRC64;

Query Match 79.5%; Score 31; DB 16; Length 578;
 Best Local Similarity 66.7%; Pred. No. 15e-02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GVNAXASL 9
 113 GVNAXASL 207

RESULT 14
 Q9RCX4 PRELIMINARY; PRT; 63 AA.
 ID Q9RCX4;
 AC Q9RCX4;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
 DE Epsw protein.
 GN Epsw.
 OS Streptococcus thermophilus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 RN NCBI_TaxID=1308;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98325267; PubMed=9657932;
 RA Bourgoin F., Guedon G., Gintz B., Decaris B.;
 RT "Characterization of a novel insertion sequence, IS1194, in
 RT Streptococcus thermophilus";
 RL Plasmid 40:44-49(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99307153; PubMed=10375631;
 RA Bourgoin F., Pluvinet A., Gintz B., Decaris B.;
 RT "Are horizontal transfers involved in the evolution of the
 RT Streptococcus thermophilus exopolysaccharide synthesis loci?";
 RL Gene 233:151-161(1999).
 DR EMBL; Z58171; CAB52220.1; -
 SQ SEQUENCE 63 AA; 7278 MW; ICCA6A7C6F2D6926 CRC64;

Query Match 76.9%; Score 30; DB 2; Length 63;
 Best Local Similarity 66.7%; Pred. No. 24;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GVNAXASL 9
 6 GVNAXASL 14

RESULT 15
 Q8YK27 PRELIMINARY; PRT; 168 AA.
 ID Q8YK27;
 AC Q8YK27;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Transposase.
 GN ALR8502.
 OS Anabaena sp. (strain PCC 7120).
 OC Cyanobacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 RN NCBI_TaxID=103690;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11758940;
 RA Karako T., Nakamura Y., Wolz C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iritsuchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120";
 RL DNA Res. 8:205-213(2001).
 DR EMBL; AP003604; BAB77421.1; -
 KM Plasmid; Complete proteome.
 SQ SEQUENCE 168 AA; 21357 MW; 014267C02480F5A8 CRC64;

Query Match 76.9%; Score 30; DB 16; Length 168;
 Best Local Similarity 66.7%; Pred. No. 77;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GVNAXASL 9
 113 GVNAXASL 207

Db |:| |||
 5 GLSATSLF 13

Search completed: January 29, 2004, 14:56:49
Job time : 27.8676 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:52:19 (Search time 11.1176 Seconds
(without alignments)
34,252 Million cell updates/sec

Title: US-10-032-950-5
Perfect score: 39
Sequence: 1 GVNAXASLF 9

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:
1: /cgm2_6/prodata/1/aa/5A.COMB.pep:*
2: /cgm2_6/prodata/1/aa/5B.COMB.pep:*
3: /cgm2_6/prodata/1/aa/6A.COMB.pep:*
4: /cgm2_6/prodata/1/aa/6B.COMB.pep:*
5: /cgm2_6/prodata/1/aa/PCITUS.COMB.pep:*
6: /cgm2_6/prodata/1/aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match Length	DB ID	Description
1	37	94.9	9 4 US-09-339-511-5	Sequence 5, Appli
2	34	87.2	9 4 US-09-339-511-1	Sequence 1, Appli
3	33	84.6	9 4 US-08-861-476C-3	Sequence 3, Appli
4	33	84.6	47 4 US-08-861-476C-6	Sequence 6, Appli
5	31	79.5	9 4 US-09-339-511-6	Sequence 6, Appli
6	30	76.9	9 4 US-09-339-511-2	Sequence 2, Appli
7	30	76.9	339 4 US-09-107-532A-4904	Sequence 4904, Ap
8	30	76.9	464 4 US-09-252-991A-27559	Sequence 27559, A
9	30	76.9	464 4 US-09-252-991A-28048	Sequence 28048, A
10	29	74.4	9 4 US-09-339-511-7	Sequence 7, Appli
11	29	74.4	332 4 US-09-882-835-2	Sequence 2, Appli
12	29	74.4	488 1 US-08-115-365-2	Sequence 2, Appli
13	29	74.4	488 1 US-08-586-897-2	Sequence 2, Appli
14	28	71.8	9 4 US-09-339-511-4	Sequence 4, Appli
15	28	71.8	205 4 US-09-252-991A-25625	Sequence 25625, A
16	28	71.8	233 4 US-09-252-991A-20791	Sequence 20791, A
17	28	71.8	1415 4 US-09-252-991A-26438	Sequence 26438, A
18	28	71.8	2154 2 US-08-841-349-4	Sequence 4, Appli
19	27	69.2	32 3 US-08-484-322-56	Sequence 56, Appli
20	27	69.2	69 4 US-09-134-001C-4106	Sequence 4106, A
21	27	69.2	217 4 US-09-252-991A-29222	Sequence 29222, A
22	27	69.2	264 4 US-09-328-352-6033	Sequence 6033, Ap
23	27	69.2	264 4 US-09-328-352-7074	Sequence 7074, Ap
24	27	69.2	237 4 US-09-830-217-18	Sequence 18, Appli
25	27	69.2	348 4 US-09-134-001C-4857	Sequence 4857, A
26	27	69.2	350 4 US-09-252-991A-33358	Sequence 33358, A
27	27	69.2	436 4 US-09-252-991A-22426	Sequence 22426, A

28	27	69.2	442 4 US-09-338-352-5748	Sequence 5748, Ap
29	27	69.2	447 4 US-09-252-991A-20563	Sequence 20563, A
30	27	69.2	469 4 US-09-198-452A-793	Sequence 793, App
31	27	69.2	488 4 US-09-134-001C-4246	Sequence 4246, Ap
32	27	69.2	501 1 US-08-499-215-4	Sequence 4, Appli
33	27	69.2	597 1 US-08-462-884A-1	Sequence 1, Appli
34	27	69.2	597 1 US-08-461-881B-1	Sequence 1, Appli
35	27	69.2	597 2 US-09-123-960-1	Sequence 1, Appli
36	27	69.2	605 1 US-08-462-884A-3	Sequence 3, Appli
37	27	69.2	605 1 US-08-461-881B-3	Sequence 3, Appli
38	27	69.2	605 2 US-09-123-960-3	Sequence 3, Appli
39	27	69.2	606 4 US-09-347-878-34	Sequence 34, Appli
40	27	69.2	804 4 US-09-252-991A-23789	Sequence 23789, A
41	27	69.2	1548 4 US-09-252-991A-22301	Sequence 22301, A
42	27	69.2	2887 3 US-08-462-467B-2	Sequence 2, Appli
43	27	69.2	2887 3 US-08-462-467B-8	Sequence 8, Appli
44	26	66.7	9 4 US-09-339-511-3	Sequence 3, Appli
45	26	66.7	35 4 US-09-205-258-436	Sequence 436, App

ALIGNMENTS

```

RESULT 1
US-09-339-511-5
Sequence 5, Application US/09339511
Patent No. 6337385
GENERAL INFORMATION:
APPLICANT: Muir, Tom
APPLICANT: Mayville, Patricia
APPLICANT: No. 6337385ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: Ji, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
FILE REFERENCE: 600-1-231n
CURRENT APPLICATION NUMBER: US/09/339, 511
CURRENT FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: 60/050,402
PRIOR FILING DATE: 1998-06-24
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 5
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-5
Query Match 94.9%; Score 37; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GVNAXASLF 9
Db 1 GVNAXASLF 9
RESULT 2
US-09-339-511-1
Sequence 1, Application US/09339511
Patent No. 6337385
GENERAL INFORMATION:
APPLICANT: Muir, Tom
APPLICANT: Mayville, Patricia
APPLICANT: No. 6337385ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: Ji, Guangyong

```


TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
TITLE OF INVENTION: INTERFERENCE
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/09/339,511
CURRENT FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 1
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: peptide
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-1

Query Match 87.2%; Score 34; DB 4; Length 9;
Best Local Similarity 88.9%; Pred. No. 2.5e+05;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVNAXSLF 9
Db 1 GVNAXSLF 9

RESULT 3
US-08-861-476C-3
Sequence 3, Application US/08861476C
Patent No. 6447786
GENERAL INFORMATION:
APPLICANT: New York University Medical Center
TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
FILE REFERENCE: 63753/7
CURRENT APPLICATION NUMBER: US/08/861,476C
CURRENT FILING DATE: 1997-05-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin version 3.0
SEQ ID NO: 3
LENGTH: 9
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-08-861-476C-3

Query Match 84.6%; Score 33; DB 4; Length 9;
Best Local Similarity 77.8%; Pred. No. 2.5e+05;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSLF 9
Db 1 GVNAXSLF 9

RESULT 4
US-08-861-476C-6
Sequence 6, Application US/08861476C
Patent No. 6447786
GENERAL INFORMATION:
APPLICANT: New York University Medical Center
TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
FILE REFERENCE: 63753/7
CURRENT APPLICATION NUMBER: US/08/861,476C
CURRENT FILING DATE: 1997-05-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin version 3.0
SEQ ID NO: 6
LENGTH: 47
TYPE: PRT
ORGANISM: Staphylococcus aureus

US-08-861-476C-6

Query Match 84.6%; Score 33; DB 4; Length 47;
Best Local Similarity 77.8%; Pred. No. 0.92;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSLF 9
Db 24 GVNAXSLF 32

RESULT 5
US-09-339-511-6
Sequence 6, Application US/09339511
Patent No. 6337385
GENERAL INFORMATION:
APPLICANT: Muir, Tom
APPLICANT: Mayville, Patricia
APPLICANT: Beavis, Ronald
APPLICANT: Ji, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
TITLE OF INVENTION: INTERFERENCE
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/09/339,511
CURRENT FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 6
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: peptide
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-6

Query Match 79.5%; Score 31; DB 4; Length 9;
Best Local Similarity 77.8%; Pred. No. 2.5e+05;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVNAXSLF 9
Db 1 GVNAXSLF 9

RESULT 6
US-09-339-511-2
Sequence 2, Application US/09339511
Patent No. 6337385
GENERAL INFORMATION:
APPLICANT: Muir, Tom
APPLICANT: Mayville, Patricia
APPLICANT: Beavis, Ronald
APPLICANT: Ji, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
TITLE OF INVENTION: INTERFERENCE
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/09/339,511
CURRENT FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 2
LENGTH: 9
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: peptide
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-2

Query Match 76.9%; Score 30; DB 4; Length 9;
Best Local Similarity 77.8%; Pred. No. 2.5e+05;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYNAXASLF 9
DB 1 GANAXSLF 9

RESULT 7
US-09-107-532A-4904
Sequence 4904, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 4804:

SEQUENCE CHARACTERISTICS:
LENGTH: 339 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...339

SEQUENCE DESCRIPTION: SEQ ID NO: 4904:

US-09-107-532A-4904

Query Match 76.9%; Score 30; DB 4; Length 339;
Best Local Similarity 75.0%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYNAXASL 8
DB 104 GINAGASL 111

RESULT 8
US-09-252-991A-27559
Sequence 27559, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 27559
LENGTH: 464
TYPE: PRT

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27559

Query Match 76.9%; Score 30; DB 4; Length 464;
Best Local Similarity 66.7%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYNAXASLF 9
DB 60 GINLPASLF 68

RESULT 9
US-09-252-991A-28048
Sequence 28048, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 28048
LENGTH: 464
TYPE: PRT

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28048

Query Match 76.9%; Score 30; DB 4; Length 464;
Best Local Similarity 66.7%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYNAXASLF 9
DB 60 GINLPASLF 68

RESULT 10
US-09-339-511-7
Sequence 7, Application US/09339511
Patent No. 6337385

GENERAL INFORMATION:
APPLICANT: Mull, Tom

APPLICANT: Mayville, Patricia
APPLICANT: No. 6337385ick, Richard P.
APPLICANT: Beaver, Ronald
APPLICANT: Ji, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
TITLE OF INVENTION: INTERFERENCE
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/09/339,511
CURRENT FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 7
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-7

Query Match 74.4%; Score 29; DB 4; Length 9;
Best Local Similarity 77.8%; Pred. No. 2.5e+05;
Matches 7; Conservative 1; Indels 0; Gaps 0;

QY 1 GYNAXASLF 9
DB 1 GYNAXSSAF 9

RESULT 11
US-09-882-835-2
Sequence 2, Application US/09882835
Patent No. 6462187
GENERAL INFORMATION:
APPLICANT: Bandaru, Rajasekhar
TITLE OF INVENTION: 22109, A NOVEL HUMAN THIOREDOXIN FAMILY
TITLE OF INVENTION: MEMBER AND USES THEREOF
FILE REFERENCE: 10448-063001
CURRENT APPLICATION NUMBER: US/09/882,835
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/211,673
PRIOR FILING DATE: 2000-06-15
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 2
LENGTH: 332
TYPE: PRT
ORGANISM: Homo sapiens
US-09-882-835-2

Query Match 74.4%; Score 29; DB 4; Length 332;
Best Local Similarity 66.7%; Pred. No. 71;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYNAXASLF 9
DB 197 GYNASTPLF 205

RESULT 12
US-08-115-365-2
Sequence 2, Application US/08115365
Patent No. 5685814
GENERAL INFORMATION:
APPLICANT: ABRAMOVITZ, MARK
APPLICANT: ADAM, MOHAMED
APPLICANT: BASTIEN, LISON
APPLICANT: GRYGORCZYK, RICHARD

APPLICANT: METTERS, KATHLEEN
APPLICANT: RUSHMORE, THOMAS H.
APPLICANT: SAWYER, NICOLE
TITLE OF INVENTION: DNA ENCODING PROSTAGLANDIN RECEPTOR EP2
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: JOHN WALLEN
STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/115,365
FILING DATE: 31-AUG-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WALLEN III, JOHN W.
REGISTRATION NUMBER: 35,403
REFERENCE/DOCKET NUMBER: 19066
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-3905
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 488 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-115-365-2

Query Match 74.4%; Score 29; DB 1; Length 488;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYNAXASL 8
DB 5 GYNSSASL 12

RESULT 13
US-08-586-897-2
Sequence 2, Application US/08586897
Patent No. 5759789
GENERAL INFORMATION:
APPLICANT: ABRAMOVITZ, MARK
APPLICANT: ADAM, MOHAMED
APPLICANT: BASTIEN, LISON
APPLICANT: GRYGORCZYK, RICHARD
APPLICANT: METTERS, KATHLEEN
APPLICANT: RUSHMORE, THOMAS H.
APPLICANT: SAWYER, NICOLE
TITLE OF INVENTION: DNA ENCODING PROSTAGLANDIN RECEPTOR EP2
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: JOHN WALLEN
STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/586,897

```

; FILING DATE: 05-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/115,365
; FILING DATE: 31-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WALLEN TIT, JOHN W.
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-3905
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 488 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-586-897-2

```

```

Query Match          74.4%; Score 29; DB 1; Length 488;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 GVNAXASL 8
      |||:||||
Db      5 GVNSSASL 12

```

```

RESULT 14
; US-09-339-511-4
; Sequence 4; Application US/09339511
; Patent No. 6337385
; GENERAL INFORMATION:
; APPLICANT: Multicom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. 6337385ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/09/339,511
; CURRENT FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
; US-09-339-511-4

```

```

Query Match          71.8%; Score 28; DB 4; Length 9;
Best Local Similarity 87.5%; Pred. No. 2.5e+05;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 VNAXASLF 9
      |||:||||
Db      2 VNAXSFLF 9

```

```

RESULT 15
; US-09-252-991A-25625
; Sequence 25625; Application US/09252991A
; Patent No. 6551795

```

```

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25625
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-25625

```

```

Query Match          71.8%; Score 28; DB 4; Length 205;
Best Local Similarity 75.0%; Pred. No. 69;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 GVNAXASL 8
      |||:||||
Db      71 GINAKASL 78

```

```

Search completed: January 29, 2004, 15:03:22
Job time : 12.1176 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:47:39 ; Search time 25.5441 Seconds
(without alignments)
73.223 Million cell updates/sec

Title: US-10-032-950-5

Sequence: 1 GYNAXASLF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PTCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PTCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	94.9	9	13	US-10-032-950-5
2	34	87.2	9	13	US-10-032-950-1
3	33	84.6	9	15	US-10-201-444-3
4	33	84.6	47	15	US-10-201-444-6
5	32	82.1	490	15	US-10-245-351-2
6	31	79.5	9	13	US-10-032-950-6
7	31	79.5	279	12	US-10-369-493-11888
8	31	79.5	800	13	US-10-369-493-14064
9	30	76.9	9	13	US-10-032-950-2
10	30	76.9	147	12	US-09-864-408A-7334
11	30	76.9	415	12	US-10-292-798-1638
12	30	76.9	432	12	US-10-017-161-2252
13	30	76.9	432	12	US-10-292-798-1898
14	29	74.4	9	13	US-10-032-950-7
15	29	74.4	332	12	US-10-145-586-46

16	29	74.4	468	12	US-10-369-493-12194	Sequence 12194, A
17	29	74.4	488	12	US-10-320-351-3	Sequence 3, Appl1
18	29	74.4	488	12	US-09-826-509-561	Sequence 561, App
19	29	74.4	488	12	US-10-295-027-1326	Sequence 1326, Ap
20	29	74.4	488	15	US-10-225-567A-296	Sequence 296, App
21	29	74.4	605	9	US-09-841-132-574	Sequence 574, App
22	29	74.4	747	10	US-09-978-295A-459	Sequence 459, App
23	29	74.4	747	10	US-09-978-418-9	Sequence 9, Appl1
24	29	74.4	747	10	US-09-978-459	Sequence 459, App
25	29	74.4	747	10	US-09-978-192A-459	Sequence 459, App
26	29	74.4	747	10	US-09-999-832A-459	Sequence 459, App
27	29	74.4	747	11	US-09-978-189-459	Sequence 459, App
28	29	74.4	747	11	US-09-978-608A-459	Sequence 459, App
29	29	74.4	747	11	US-09-978-585A-459	Sequence 459, App
30	29	74.4	747	11	US-09-978-191A-459	Sequence 459, App
31	29	74.4	747	11	US-09-978-403A-459	Sequence 459, App
32	29	74.4	747	11	US-09-978-564A-459	Sequence 459, App
33	29	74.4	747	11	US-09-999-833A-459	Sequence 459, App
34	29	74.4	747	11	US-09-981-915A-459	Sequence 459, App
35	29	74.4	747	11	US-09-978-824-459	Sequence 459, App
36	29	74.4	747	11	US-09-978-585A-459	Sequence 459, App
37	29	74.4	747	11	US-09-978-423A-459	Sequence 459, App
38	29	74.4	747	11	US-09-978-193A-459	Sequence 459, App
39	29	74.4	747	11	US-09-999-830A-459	Sequence 459, App
40	29	74.4	747	11	US-09-978-157A-459	Sequence 459, App
41	29	74.4	747	11	US-09-978-187A-459	Sequence 459, App
42	29	74.4	747	11	US-09-978-643A-459	Sequence 459, App
43	29	74.4	747	12	US-09-978-375A-459	Sequence 459, App
44	29	74.4	747	12	US-09-978-188A-459	Sequence 459, App
45	29	74.4	747	12	US-09-978-298A-459	Sequence 459, App

ALIGNMENTS

RESULT 1
US-10-032-950-5
; Sequence 5, Application US/10032950
; Publication No. US20020077453A1
GENERAL INFORMATION:
; APPLICANT: Walt, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard F.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 5
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-5

Query Match 94.9% ; Score 37; DB 13; Length 9;
Best Local Similarity 100.0% ; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 1 GYNAXASLF 9
| | | | |
Db 1 GYNAXASLF 9

```
RESULT 2
US-10-032-950-1
; Sequence 1, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Muir, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-1

Query Match
Best Local Similarity 87.2%; Score 34; DB 13; Length 9;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 1 GVNAXASLF 9
Db 1 GVNAXASLF 9

RESULT 3
US-10-201-444-3
; Sequence 3, Application US/10201444
; Publication No. US20030078378A1
; GENERAL INFORMATION:
; APPLICANT: New York University Medical Center
; TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
; FILE REFERENCE: 63753/7
; CURRENT APPLICATION NUMBER: US/10/201,444
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/08/861,476
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-201-444-3

Query Match
Best Local Similarity 84.6%; Score 33; DB 15; Length 9;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Cy 1 GVNAXASLF 9
Db 1 GVNAXASLF 9

RESULT 4
US-10-201-444-6
```

```
; Sequence 6, Application US/10201444
; Publication No. US20030078378A1
; GENERAL INFORMATION:
; APPLICANT: New York University Medical Center
; TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
; FILE REFERENCE: 63753/7
; CURRENT APPLICATION NUMBER: US/10/201,444
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/08/861,476
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-201-444-6

Query Match
Best Local Similarity 84.6%; Score 33; DB 15; Length 47;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Cy 1 GVNAXASLF 9
Db 24 GVNAXASLF 32

RESULT 5
US-10-243-351-2
; Sequence 2, Application US/10243351
; Publication No. US20030059890A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc.
; APPLICANT: Castleberry, Tessa A.
; APPLICANT: Lu, Bihong
; APPLICANT: Owen, Thomas A.
; APPLICANT: Smock, Steven
; TITLE OF INVENTION: DNA ENCODING THE CHIMPANZEE PROSTAGLANDIN E2 RECEPTOR EP4 SUBT
; FILE REFERENCE: PCT/US2002/007823
; CURRENT APPLICATION NUMBER: US/10/243,351
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 60/322,915
; PRIOR FILING DATE: 2001-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 490
; TYPE: PRT
; ORGANISM: chimpanzee
US-10-243-351-2

Query Match
Best Local Similarity 82.1%; Score 32; DB 15; Length 490;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 GVNAXASL 8
Db 5 GVNAXASL 12

RESULT 6
US-10-032-950-6
; Sequence 6, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Muir, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
```

```
/ CURRENT APPLICATION NUMBER: US/10/032,950
/ CURRENT FILING DATE: 2001-12-27
/ PRIOR APPLICATION NUMBER: 60/090,402
/ PRIOR FILING DATE: 1998-06-24
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: Patent Ver. 2.0
/ SEQ ID NO 6
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic
/ OTHER INFORMATION: peptide
/ NAME/KEY: VARIANT
/ LOCATION: (5)
/ OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-6
```

```
Query Match      79.5%; Score 31; DB 13; Length 9;
Best Local Similarity 77.8%; Pred. No. 7e+05;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GVNAXASLF 9
      |||||
Db      1 GVNAXASLF 9

RESULT 7
US-10-369-493-11888
/ Sequence 11888, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 11888
/ LENGTH: 279
/ TYPE: PRT
/ ORGANISM: Mesorhizobium loti
US-10-369-493-11888
```

```
Query Match      79.5%; Score 31; DB 12; Length 279;
Best Local Similarity 87.5%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GVNAXASL 8
      |||||
Db      139 GVNAXASL 146
```

```
RESULT 8
US-10-369-493-14064
/ Sequence 14064, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
```

```
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 14064
/ LENGTH: 800
/ TYPE: PRT
/ ORGANISM: Pseudomonas fluorescens
US-10-369-493-14064
```

```
Query Match      79.5%; Score 31; DB 12; Length 800;
Best Local Similarity 77.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 GVNAXASLF 9
      |||||
Db      133 GVTALASLF 141
```

```
RESULT 9
US-10-032-950-2
/ Sequence 2, Application US/10032950
/ Publication No. US20020077453A1
/ GENERAL INFORMATION:
/ APPLICANT: Muir, Tom
/ APPLICANT: Mayville, Patricia
/ APPLICANT: No. US20020077453A1ick, Richard P.
/ APPLICANT: Beavis, Ronald
/ APPLICANT: Ji, Guangyong
/ TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
/ FILE REFERENCE: 600-1-231N
/ CURRENT APPLICATION NUMBER: US/10/032,950
/ CURRENT FILING DATE: 2001-12-27
/ PRIOR APPLICATION NUMBER: 60/090,402
/ PRIOR FILING DATE: 1998-06-24
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: Patent Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic
/ NAME/KEY: VARIANT
/ LOCATION: (5)
/ OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-2
```

```
Query Match      76.9%; Score 30; DB 13; Length 9;
Best Local Similarity 77.8%; Pred. No. 7e+05;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 GVNAXASLF 9
      |||||
Db      1 GVNAXSLSLF 9
```

```
RESULT 10
US-09-864-408A-7334
/ Sequence 7334, Application US/09864408A
/ Publication No. US20040009474A1
/ GENERAL INFORMATION:
/ APPLICANT: Leach, Martin D.
/ APPLICANT: Shinkets, Richard A.
/ TITLE OF INVENTION: NO. US20040009474A1el Human Polynucleotides and Polypeptides E:
/ FILE REFERENCE: 21402-012
/ CURRENT APPLICATION NUMBER: US/09/864,408A
/ CURRENT FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: 60/206,690
/ PRIOR FILING DATE: 2000-05-24
```

NUMBER OF SEQ ID NOS: 9068
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7334
LENGTH: 147
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1)
OTHER INFORMATION: wherein Xaa may be any naturally occurring amino acid
US-09-864-408A-7334

Query Match
Best Local Similarity 76.9%; Score 30; DB 12; Length 147;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNXASL 9
Db 41 GINLPASL 49

RESULT 11
US-10-292-798-1638
Sequence 1638, Application US/10292798
Publication No. US2003023583A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURAYANI, HIROUKI
TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
FILE REFERENCE: 084335/166
CURRENT APPLICATION NUMBER: US/10/292,798
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 10/017,161
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: JP 2001-246789
NUMBER OF SEQ ID NOS: 2070
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1638
LENGTH: 415
TYPE: PRT
ORGANISM: Homo sapiens
US-10-292-798-1638

Query Match
Best Local Similarity 76.9%; Score 30; DB 12; Length 415;
Best Local Similarity 66.7%; Pred. No. 94;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNXASL 9
Db 401 GWHAAASL 409

RESULT 12
US-10-017-161-2252
Sequence 2252, Application US/10017161
Publication No. US2003014368A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURAYANI, HIROUKI
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 084335/0152
CURRENT APPLICATION NUMBER: US/10/017,161
CURRENT FILING DATE: 2002-12-18
PRIOR APPLICATION NUMBER: JP 2001/246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2430
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2252

LENGTH: 432
TYPE: PRT
ORGANISM: Homo sapiens
US-10-017-161-2252

Query Match
Best Local Similarity 76.9%; Score 30; DB 12; Length 432;
Best Local Similarity 66.7%; Pred. No. 98;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNXASL 9
Db 418 GWHAAASL 426

RESULT 13
US-10-292-798-1898
Sequence 1898, Application US/10292798
Publication No. US2003023583A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURAYANI, HIROUKI
TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
FILE REFERENCE: 084335/166
CURRENT APPLICATION NUMBER: US/10/292,798
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 10/017,161
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: JP 2001-246789
NUMBER OF SEQ ID NOS: 2070
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1898
LENGTH: 432
TYPE: PRT
ORGANISM: Homo sapiens
US-10-292-798-1898

Query Match
Best Local Similarity 76.9%; Score 30; DB 12; Length 432;
Best Local Similarity 66.7%; Pred. No. 98;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNXASL 9
Db 418 GWHAAASL 426

RESULT 14
US-10-032-950-7
Sequence 7, Application US/10032950
Publication No. US20020077453A1
GENERAL INFORMATION:
APPLICANT: Muir, Tom
APPLICANT: Mayville, Patricia
APPLICANT: No. US20020077453A1, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: Li, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
TITLE OF INVENTION: INTERFERENCE
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/10/032,950
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic

OTHER INFORMATION: peptide
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-7

Query Match 74.4%; Score 29; DB 13; Length 9;
Best Local Similarity 77.8%; Pred. No. 7e+05;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GVNAXASLF 9
Db 1 GVNAXSAP 9

RESULT 15
US-10-145-586-46
Sequence 46, Application US/10145586
Publication No. US20030138890A1
GENERAL INFORMATION:
APPLICANT: Alexandra Glucksmann, Maria
APPLICANT: Silos-Santiago, Imaculada
APPLICANT: M. Galvin, Katherine
APPLICANT: Welch, Nadine
APPLICANT: Curtis, Rory A.J.
APPLICANT: Bandaru, Rajasekhar
APPLICANT: Kapeller-Libermann, Rosana
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTOR FAMILY MEMBERS,
TITLE OF INVENTION: HUMAN THIOREDOXIN FAMILY MEMBERS, HUMAN LEUCINE-RICH
TITLE OF INVENTION: REPEAT FAMILY MEMBERS, AND HUMAN RINGFINGER FAMILY MEMBER
FILE REFERENCE: 10449-188001
CURRENT APPLICATION NUMBER: US/10/145,586
CURRENT FILING DATE: 2002-05-14
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 95
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 46
LENGTH: 332
TYPE: PRT
ORGANISM: Homo sapiens
US-10-145-586-46

Query Match 74.4%; Score 29; DB 12; Length 332;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 GVNAXASLF 9
Db 197 GVNAXSAP 205

Search completed: January 29, 2004, 15:01:46
Job time : 26.5441 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:44:54 ; Search time 35.0735 Seconds
(without alignments)
40.730 Million cell updates/sec

Title: US-10-032-950-6

Perfect score: 39

Sequence: 1 GVNXSAIF 9

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_19JUN03: *
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	94.9	9	21	AA67856
2	37	94.9	9	23	ABP5345
3	37	94.9	9	23	AA50904
4	35	89.7	9	21	AA67859
5	35	89.7	9	21	AA67860
6	35	89.7	9	23	AA50907
7	35	89.7	9	23	AA51002
8	35	89.7	9	23	AA51003
9	35	89.7	9	23	AA51004

10	35	89.7	9	23	AB507160	Peptide-mediated q
11	34	87.2	9	21	AA67851	S. aureus peptide
12	34	87.2	9	21	AA67861	Staphylococcus aur
13	34	87.2	9	23	ABP5345	Cyclic peptide SEQ
14	34	87.2	9	23	AA50899	Agpd-autoinducing
15	34	87.2	9	23	AA50908	Protected peptide
16	34	87.2	9	23	AA51005	Agpd2 lactam cycli
17	34	87.2	9	23	AB507161	Peptide-mediated q
18	33	84.6	9	19	AAW38323	Transcription inh
19	33	84.6	9	23	AA51001	Agpd2 thiololacton
20	33	84.6	9	23	AB507159	Peptide-mediated q
21	33	84.6	9	24	AB584631	S. aureus RN6607 A
22	33	84.6	9	24	AB584634	S. aureus peptide
23	31	79.5	9	21	AA67855	Cyclic peptide SEQ
24	31	79.5	9	23	ABP53544	Agpd-autoinducing
25	31	79.5	9	23	AAW50903	S. aureus peptide
26	30	76.9	9	21	AA67852	Cyclic peptide SEQ
27	30	76.9	9	23	ABP53541	Agpd-autoinducing
28	30	76.9	9	23	AAW50900	Novel human diagno
29	30	76.9	103	22	AB524179	Novel human diagno
30	30	76.9	338	24	ABP50927	N. gonorrhoeae ami
31	30	76.9	389	21	AA68973	Cp61H protein whic
32	30	76.9	549	19	AAW98659	H. pylori GHP 467
33	29	74.4	9	21	AA67857	S. aureus peptide
34	29	74.4	9	23	ABP53546	Cyclic peptide SEQ
35	29	74.4	9	23	AAW50905	Agpd-autoinducing
36	29	74.4	96	22	AB510434	Novel human diagno
37	29	74.4	255	23	AB549661	Listeria monocytog
38	29	74.4	282	23	ABP26477	Streptococcus poly
39	29	74.4	397	20	AAV37161	Amino acid sequenc
40	29	74.4	605	23	AB594351	Chlamydia trachoma
41	29	74.4	895	23	AB592623	Heblichidially activ
42	29	74.4	1225	23	AB566083	Bitidbacterium lo
43	29	74.4	1350	22	AB509297	Novel human diagno
44	28	71.8	9	21	AA67854	S. aureus peptide
45	28	71.8	9	23	ABP53543	Cyclic peptide SEQ

ALIGNMENTS

RESULT 1	AA67856	standard; peptide; 9 AA.
ID	AA67856;	
AC	AA67856;	
DT	25-APR-2000	(first entry)
DE	S. aureus peptide #6 used for bacterial interference.	
XX	Staphylococcus aureus infection; cyclic peptide; Agpd; agr response;	
KW	virulence factor; treatment.	
XX	Staphylococcus aureus.	
OS	Staphylococcus aureus.	
XX		
FH	Key	Location/Qualifiers
FT	Modified-site	1
FT	/note= "N-terminal residue forms bond with C-terminal	
FT	residue to form a cyclic peptide"	
FT	Misc-difference 5	/label= Unknown
FT		
PN	WO967286-A2.	
XX		
PD	29-DEC-1999.	
XX		
PF	24-UN-1999;	99WO-US14562.
XX		
PR	24-UN-1998;	98US-0103438.
XX		
PA	(UVRQ) UNIV ROCKEFELLER.	
PA	(UVRV) UNIV NEW YORK STATE.	

XX Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
 XX WPI; 2000-147202/13.
 XX New cyclic peptides for treating infections with *Staphylococcus aureus*
 PT
 XX
 XX Claim 9; Page 26; 37pp; English.
 CC This sequence represents a cyclic peptide derived from the *Staphylococcus*
 CC *aureus* AgrD peptide. The invention relates to AgrD derived peptides,
 CC a composition containing a peptide and a carrier, and a method for the
 CC production of the cyclic peptides. The peptide inhibits the agr response,
 CC which is normally associated with the release of virulence factors of
 CC *Staphylococcus aureus*. An AgrD peptide is produced by *S. aureus* that
 CC activates the agr response in strains of a single group, but interferes
 CC with this response in strains of different groups. The peptides and
 CC composition containing them can be used to treat infections by *S. aureus*.
 CC
 SQ Sequence 9 AA;
 Query Match 94.9%; Score 37; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVNAXSALF 9
 DB 1 GVNAXSALF 9
 RESULT 2
 ABP53545
 ID ABP53545 standard; peptide; 9 AA.
 AC ABP53545;
 XX
 DT 13-DEC-2002 (first entry)
 DE Cyclic peptide SEQ ID NO.6.
 XX
 KM Cyclic peptide; *Staphylococcus aureus*; infection; antibacterial;
 KM agr response inhibitor.
 XX
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 5 /note= "any amino acid"
 XX
 PN US2002077453-A1.
 PD 20-JUN-2002.
 PF 27-DEC-2001; 2001US-0032950.
 XX
 PR 24-JUN-1998; 98US-090402P.
 PR 24-JUN-1999; 99US-0339511.
 XX
 PA (MUIR/) MUIR T W.
 PA (MAYV/) MAYVILLE P.
 PA (NOVI/) NOVICK R P.
 PA (BEAV/) BEAVIS R.
 PA (JIGG/) JI G.
 XX
 PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
 XX WPI; 2002-661366/73.
 DR
 XX New cyclic peptides, useful for treating *Staphylococcus aureus*
 PT infections
 XX
 PS Claim 9; Page 10; 18pp; English.

XX ABP53540 to ABP53547 represent cyclic peptides (1) from the present
 CC invention. The present invention also describes a method for treating
 CC *Staphylococcus aureus* infection comprising the administration of a
 CC composition comprising (1). (1) has antibacterial activity, and can be
 CC used as an agr gene response inhibitor. The peptides are useful for
 CC treating *S. aureus* infections.
 CC
 SQ Sequence 9 AA;
 Query Match 94.9%; Score 37; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVNAXSALF 9
 DB 1 GVNAXSALF 9
 RESULT 3
 AAM50904
 ID AAM50904 standard; Peptide; 9 AA.
 AC AAM50904;
 XX
 DT 08-MAY-2002 (first entry)
 DE AgrD-autoinducing cyclic peptide, inhibitor of agr response.
 XX
 KM *Staphylococcus aureus*; AgrD; agr response; inhibitor; antibiotic;
 KM antibacterial; infection; therapy; cyclic.
 XX
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 5 /note= "any amino acid"
 FT Misc-difference 5 /note= "note linked to residue 9 to form cyclic
 FT peptide"
 FT Misc-difference 9 /note= "note linked to residue 5 to form cyclic
 FT peptide"
 XX
 PN US6373385-B1.
 PD 08-JUN-2002.
 PF 24-JUN-1999; 99US-0339511.
 XX
 PR 24-JUN-1998; 98US-090402P.
 XX
 PA (UYRQ) UNIV ROCKEFELLER.
 PA (UYNY) UNIV NEW YORK STATE.
 XX
 PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
 XX WPI; 2002-170774/22.
 DR
 XX Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
 PT interference and for treating *Staphylococcus aureus* infection in a
 PT subject
 XX
 PS Claim 7; Column 19; 18pp; English.
 CC The present sequence is that of a novel synthetic cyclic peptide
 CC of the invention that is capable of inhibiting the agr response of
 CC *Staphylococcus aureus*. It is an AgrD-autoinducing peptide, where
 CC AgrD is a secreted agr-encoded peptide and where the agr locus
 CC controls the synthesis of virulence factor and other extracellular
 CC proteins responsible for pathogenicity in *S. aureus*. Preferred
 CC peptides may have the sequence NH₂-X(n)-Z-X(y)-COOH, with a
 CC cyclic bond between the Z residue and COOH other than a thioester

CC bond, where X is an amino acid, an amino acid analogue, a
 CC peptidomimetic or non-amide isostere, Z is a synthetic or a
 CC biosynthetic amino acid, n is 0-10 and y is 1-10. The cyclic bond
 CC is especially a lactam or lactone bond. The thiololactone
 CC structure within native AgrD peptides is required for activation
 CC of the agr response. Elimination of the thiol ester component of
 CC the cyclic ring structure can destroy agr response activating
 CC activity while preserving and enhancing inhibitory activity. A
 CC claimed method of preparing a cyclic peptide involves: assembling
 CC a linear peptide chain on to a solid phase resin support;
 CC deprotecting the resulting protected assembled peptide; treating the
 CC deprotected peptide with neutral buffer for a time sufficient to
 CC form the cyclic peptide and cleave the peptide from the support;
 CC and recovering the cyclic peptide. The peptide is useful for
 CC bacterial interference, especially for the treatment of *S. aureus*
 CC infection.

CC Sequence 9 AA;

Query Match 94.9%; Score 37; DB 23; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
 Db 1 GVNAXSALF 9

RESULT 4
 AA67859
 ID AA67859 standard; peptide; 9 AA.

XX AA67859;

DT 25-APR-2000 (first entry)

XX Staphylococcus aureus AgrDII derived peptide sequence.

XX Staphylococcus aureus infection; treatment; AgrD; agr response;

XX virulence factor.

XX Staphylococcus aureus.

XX WO967286-A2.

XX 29-DEC-1999.

XX 24-JUN-1999; 99WO-US14562.

XX 24-JUN-1998; 98US-0103438.

XX (UYRQ) UNIV ROCKEFELLER.

XX (UYNY) UNIV NEW YORK STATE.

XX Muir TW, Mayville P, Novick RP, Ji G, Beavis R;

XX WPI; 2000-147202/13.

XX New cyclic peptides for treating infections with Staphylococcus aureus

XX Examples; Page 22; 37pp; English.

CC This sequence represents the Staphylococcus aureus AgrDII derived
 CC peptide. The invention relates to AgrD derived cyclic peptides, a
 CC composition containing a peptide and a carrier, and a method for the
 CC production of the cyclic peptides. The peptide inhibits the agr response,
 CC which is normally associated with the release of virulence factors of
 CC Staphylococcus aureus. An AgrD peptide is produced by *S. aureus* that
 CC activates the agr response in strains of a single group, but interferes
 CC with this response in strains of different groups. The peptides and
 CC composition containing them can be used to treat infections by *S. aureus*.

SQL Sequence 9 AA;

Query Match 89.7%; Score 35; DB 21; Length 9;
 Best Local Similarity 77.8%; Pred. No. 9.3e+05;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
 Db 1 GVNAXSALF 9

RESULT 5
 AA67860
 ID AA67860 standard; peptide; 9 AA.

XX AA67860;

DT 25-APR-2000 (first entry)

XX Staphylococcus aureus AgrDII derived peptide sequence.

XX Staphylococcus aureus infection; AgrD; agr response; treatment;

XX virulence factor.

XX Staphylococcus aureus.

XX WO967286-A2.

XX 29-DEC-1999.

XX 24-JUN-1999; 99WO-US14562.

XX 24-JUN-1998; 98US-0103438.

XX (UYRQ) UNIV ROCKEFELLER.

XX (UYNY) UNIV NEW YORK STATE.

XX Muir TW, Mayville P, Novick RP, Ji G, Beavis R;

XX WPI; 2000-147202/13.

XX New cyclic peptides for treating infections with Staphylococcus aureus

XX Examples; Page 22; 37pp; English.

CC This sequence represents the Staphylococcus aureus AgrDII derived
 CC peptide. The invention relates to AgrD derived cyclic peptides, a
 CC composition containing a peptide and a carrier, and a method for the
 CC production of the cyclic peptides. The peptide inhibits the agr response,
 CC which is normally associated with the release of virulence factors of
 CC Staphylococcus aureus. An AgrD peptide is produced by *S. aureus* that
 CC activates the agr response in strains of a single group, but interferes
 CC with this response in strains of different groups. The peptides and
 CC composition containing them can be used to treat infections by *S. aureus*.

SQL Sequence 9 AA;

Query Match 89.7%; Score 35; DB 21; Length 9;
 Best Local Similarity 77.8%; Pred. No. 9.3e+05;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
 Db 1 GVNAXSALF 9

RESULT 6
 AAM50907
 ID AAM50907 standard; Peptide; 9 AA.

XX AAM50907;

XX	08-MAY-2002 (first entry)
DE	Protected peptide used in cyclic peptide production.
XN	Staphylococcus aureus; AgrD; agr response; inhibitor; antibiotic;
KM	antibacterial; infection; therapy; cyclic.
XX	Synthetic.
XX	
FH	Key
FT	Modified-site
FT	1 Location/Qualifiers
FT	/note= "Z-Gly"
FT	5 Modified-site
FT	/note= "Ser(tBu)"
FT	6 Modified-site
FT	/note= "Ser(BzL)"
FT	7 Modified-site
FT	/note= "Ser(BzL)"
FT	Misc-difference
FT	5 /note= "note linked to residue 9 to form cyclic peptide"
FT	Misc-difference
FT	9 /note= "note linked to residue 5 to form cyclic peptide"
PN	US6337385-BI.
PD	08-JAN-2002.
XX	
XX	24-JUN-1999; 99US-0339511.
XX	
PR	24-JUN-1998; 98US-090402P.
PA	(UYRQ) UNIV ROCKEFELLER.
PA	(UYNV) UNIV NEW YORK STATE.
PI	Muir TM, Mayville P, Novick RP, Beavis R, Ji G;
XX	
DR	WPI; 2002-170774/22.
XX	
PT	Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial interference and for treating Staphylococcus aureus infection in a
PT	subject -
XX	
PS	Disclosure; Column 14; 18pp; English.
XX	
CC	The present sequence is that of a protected peptide used in an
CC	example of the preparation of novel synthetic cyclic peptides of
CC	the invention (see AAM50899-906). The peptide corresponds to the
CC	Staphylococcus aureus AgrDI sequence with a Cys5 to Ser mutation
CC	(lactone). It was synthesised on a Wang-resin using an Fmoc
CC	N-alpha protection strategy. Following chain assembly, the peptide
CC	was cleaved from the support and the Ser-5 residue deprotected by
CC	treatment with a trifluoroacetic acid/anisole:water mixture
CC	(90:5:5) for 4 hr. The partially protected peptide-alpha
CC	carboxylates were then dissolved in DMF and treated with PyBOP
CC	and a catalytic amount of dimethylaminopyridine. Cyclization was
CC	complete after 2 hr. The remaining protecting groups were then
CC	removed by treatment with HF and the peptide purified by HPLC.
CC	The cyclic peptide is capable of inhibiting the agr response of
CC	Staphylococcus aureus. The thiolactone structure within native
CC	AgrD peptides is required for activation of this response.
CC	Replacement of the thio ester component of the cyclic ring
CC	structure with a lactone (as in the present case) or a lactam can
CC	enhance agr response activating activity while preserving and
CC	enhancing inhibitory activity. The cyclic peptides are useful for
CC	bacterial interference, especially for the treatment of S. aureus
CC	infection.
XX	
SO	Sequence 9 AA;
Query Match	69.7%; Score 35; DB 23; Length 9;
Best Local Similarity	77.8%; Pred. No. 9.3e+05;

Matches	7;	Conservative	1;	Mismatches	1;	Indels	0;	Gaps	0;
---------	----	--------------	----	------------	----	--------	----	------	----

QY	1 GVNAXSALF 9								
	:								
Db	1 GVNAXSALP 9								


```

RESULT 7
AAM51002
ID AAM51002 standard; Peptide; 9 AA.
XX
AC AAM51002;
XX
DT 08-MAY-2002 (first entry)
XX
DE AgRD2 linear thioester peptide.
XX
KM AgRD2; agr response; inhibitor; antibiotic; antibacterial;
infection; therapy.
XX
OS Staphylococcus aureus.
XX
SY Synthetic.
XX
Key Modified-site 9 Location/Qualifiers
FT FT /note= "C-terminal thioester"
XX
PN US6337385-B1.
XX
PD 08-JAN-2002.
XX
PF 24-JUN-1999; 99US-0339511.
XX
PR 24-JUN-1998; 98US-090402P.
XX
PA (UYRO ) UNIV ROCKEFELLER.
PA (UTNY ) UNIV NEW YORK STATE.
PI
PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
DR WPI; 2002-170774/22.
XX
PT Novel synthetic, cyclic AgRD-autoinducing peptide for bacterial
PT interference and for treating Staphylococcus aureus infection in a
PT subject -
XX
XX Example 1; Column 9; 18bp; English.
XX
CC The present sequence is that of a novel synthetic AgRD2 linear
CC thioester peptide. The peptide is derived from the cyclic AgRD2
CC peptide of Staphylococcus aureus group II. AgRD2 is a secreted
CC agr-encoded peptide, where the agr locus controls the synthesis of
CC virulence factor and other extracellular proteins responsible for
CC pathogenicity in S. aureus. The biological activity of the
CC synthetic peptide was assayed using cultured S. aureus strains
CC containing a beta-lactamase reporter gene fused to the agr3
CC promoter. This allowed activation or inhibition of the agr
CC response to be monitored spectrophotometrically. Unlike an AgRD2
CC thiolactone cyclic peptide (see AAM51001), the present peptide
CC was unable to either activate or inhibit the agr response, even
CC when added to cultured cells at uM concentrations. The invention
CC provides claimed cyclic peptides (see AAM5089-96 and AAM50999)
CC and methods for preparing them. The cyclic peptides are useful for
CC bacterial interference, especially for the treatment of S. aureus
CC infection.
XX
SQ Sequence 9 AA;

```


Query Match	89.7%	Score 35;	DB 23;	Length 9;
Best Local Similarity	77.8%;	Pred No.	9.3e-05;	
Matches	7;	Conservative	1;	Mismatches 1; Indels 0; Gaps 0;
1 GVNAXSALF 9				

Db 1 GVNASSSLF 9

RESULT 8

AAMS1003 standard; Peptide; 9 AA.

AC AAMS1003;

DT 08-MAY-2002 (first entry)

DE Agrd2 linear free acid peptide.

KW Agrd2; agr response; inhibitor; antibiotic; antibacterial;

OS Staphylococcus aureus.

PN US6337385-B1.

PD 08-JAN-2002.

PF 24-JUN-1999; 99US-0339511.

PR 24-JUN-1998; 98US-090402P.

PA (UYRQ) UNIV ROCKEFELLER.

PA (UYNV) UNIV NEW YORK STATE.

PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;

DR WPI; 2002-170774/22.

PT Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial

PT interference and for treating Staphylococcus aureus infection in a

PT subject -

PS Example 1; Column 9; 18pp; English.

XX The present sequence is that of a novel synthetic AgrD2 linear

XX free acid peptide. The peptide is derived from the cyclic AgrD2

XX peptide of Staphylococcus aureus group II. AgrD2 is a secreted

XX agr-encoded peptide, where the agr locus controls the synthesis of

XX virulence factor and other extracellular proteins responsible for

XX pathogenicity in S. aureus. The biological activity of the

XX synthetic peptide was assayed using cultured S. aureus strains

XX containing a beta-lactamase reporter gene fused to the agr3

XX promoter. This allowed activation or inhibition of the agr

XX response to be monitored spectrophotometrically. Unlike an AgrD2

XX thiolactone cyclic peptide (see AAMS1001), the present peptide

XX was unable to either activate or inhibit the agr response, even

XX when added to cultured cells at 10⁶ concentrations. The invention

XX provides claimed cyclic peptides (see AAMS0899-906 and AAMS0999)

XX and methods for preparing them. The cyclic peptides are useful for

XX bacterial interference, especially for the treatment of S. aureus

XX infection.

XX AAMS1004;

XX 08-MAY-2002 (first entry)

DE Agrd2 lactone cyclic peptide.

KW Agrd2; agr response; inhibitor; antibiotic; antibacterial;

OS Staphylococcus aureus.

PN US6337385-B1.

PD 08-JAN-2002.

PF 24-JUN-1999; 99US-0339511.

PR 24-JUN-1998; 98US-090402P.

PA (UYRQ) UNIV ROCKEFELLER.

PA (UYNV) UNIV NEW YORK STATE.

PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;

DR WPI; 2002-170774/22.

PT Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial

PT interference and for treating Staphylococcus aureus infection in a

PT subject -

PS Example 1; Column 9; 18pp; English.

XX The present sequence is that of a novel synthetic AgrD2

XX lactone cyclic peptide in which residue 5 of the peptide is

XX linked to residue 9 via a lactone bond. The peptide is derived

XX from an AgrD2 peptide of Staphylococcus aureus group II. AgrD2 is

XX a secreted agr-encoded peptide, where the agr locus controls the

XX synthesis of virulence factor and other extracellular proteins

XX responsible for pathogenicity in S. aureus. The biological

XX activity of the synthetic peptide was assayed using cultured S.

XX aureus strains containing a beta-lactamase reporter gene fused to

XX the agr3 promoter. This allowed activation or inhibition of the

XX agr response to be monitored spectrophotometrically. The lactone

XX AgrD2 peptide inhibited the agr response in group I S. aureus

XX strains without activating the agr response in group II or III

XX strains. The invention provides claimed cyclic peptides (see

XX AAMS0899-906 and AAMS0999) and methods for preparing them,

XX especially peptides where the cyclic bond is a lactam or lactone

XX bond. The cyclic peptides are useful for bacterial interference,

XX especially for the treatment of S. aureus infection.

SQL Sequence 9 AA;

Query Match 89.7%; Score 35; DB 23; Length 9;

Best Local Similarity 77.8%; Pred. No. 9.3e+05;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSLF 9

Db 1 GVNASSSLF 9

RESULT 10

XX	ABB07160	standard; peptide; 9 AA.
ID	ABB07160	standard; peptide; 9 AA.
AC	ABB07160;	
XX		
DT	13-MAR-2002	(first entry)
XX		
DE	Peptide-mediated quorum sensing inhibitor peptide cyclo-XII.	
XX		
KM	Autoinducer-2; AI-2; antibiotic; antibacterial; dermatological;	
XX	vulnerable; pheromone; agr system; accessory gene regulator; cyclic.	
OS	Synthetic.	
XX		
PN	WO200185664-A2.	
PD	15-NOV-2001.	
PF	10-MAY-2001; 2001WO-US15221.	
XX		
PR	10-MAY-2000; 2000US-203000P.	
PR	07-DEC-2000; 2000US-254398P.	
XX		
PA	(UTPR-) UNIT PRINCETON.	
PA	(QOR-) QOREX PHARM INC.	
PA	(UYTE-) UNIT TECHNOLOGIES INT INC.	
PI	Bassler BL, Dammel CS, Schauder S, Shokat K, Stein J, Surette MG;	
DR	WPI; 2002-075235/10.	
XX		
PT	Use of autoinducer-2 agonists or antagonists for regulating activity of	
PT	autoinducer-2 receptor, regulating bacterial growth and pathogenesis,	
PT	also antibiotic compositions -	
XX		
PS	Disclosure; Page 33; 134pp; English.	
CC	The invention relates to the use of autoinducer-2 (AI-2) agonists or	
CC	antagonists for regulating activity of autoinducer-2 receptor, regulating	
CC	bacterial growth and pathogenesis. Synergistic antibiotic compositions	
CC	comprising inhibitors of the quorum-sensing pathway of a microorganism	
CC	are also provided. Methods using such AI-2 analogues are useful for	
CC	treating pathogen-associated disease states. The compounds and antibiotic	
CC	compositions can be used to inhibit bacterial cell growth and/or biofilm	
CC	formation on a medical device, particularly for promoting growth of skin	
CC	graft replacements used in the treatment of burns and ulcers. They may	
CC	also be used to aid wound repair, and to inhibit bacterial cell growth	
CC	and biofilm formation in or on products or devices used for personal	
CC	hygiene. The present sequence represents a inhibitor of peptide-mediated	
CC	quorum sensing.	
CC		
XX		
SO	Sequence 9 AA;	
Query Match	89.7%; Score 35; DB 23; Length 9;	
Best Local Similarity	77.8%; Pred. No. 9, 3e+05;	
Matches	7; Conservative 1; Mismatches 1; Indels 0; Gaps 0	
OY	1 GVNXSLF 9	
	:	
DB	1 GVNXSSLF 9	
RESULT 11		
AAV67851		
ID	AAV67851 standard; peptide; 9 AA.	
XX		
AC	AAV67851;	
XX		
DT	25-APR-2000 (first entry)	
XX		
DE	S. aureus peptide #1 used for bacterial interference.	
XX		
KM	Staphylococcus aureus infection; cyclic peptide; AgrD; agr response;	

KM	virulence factor; treatment.
XX	
OS	Staphylococcus aureus.
XX	
FH	Key
FT	Misc-difference 5 Location/Qualifiers
PX	/label= Unknown
PN	
XX	
MO9967286-A2.	
PD	
29-DEC-1999.	
PF	24-JUN-1999; 99WO-US14562.
XX	
PR	24-JUN-1998; 98US-0103438.
XX	
PA	(UVRQ) UNIV ROCKEFELLER.
PI	(UYNY) UNIV NEW YORK STATE.
PJ	
PI	Muir TW, Mayville P, Novick RP, Ji G, Beavis R,
XX	
DR	WPI; 2000-147202/13.
PX	
PT	New cyclic peptides for treating infections with Staphylococcus aureus
PS	-
XX	
PS	Claim 9; Page 26; 37pp; English.
XX	
CC	This sequence represents a cyclic peptide derived from the Staphylococcus
CC	aureus Agrd peptide. The invention relates to Agrd derived peptides,
CC	a composition containing a peptide and a carrier, and a method for the
CC	production of the cyclic peptides. The peptide inhibits the agr response,
CC	which is normally associated with the release of virulence factors of
CC	Staphylococcus aureus. An Agrd peptide is produced by S. aureus that
CC	activates the agr response in strains of a single group, but interferes
CC	with this response in strains of different groups. The peptides and
CC	composition containing them can be used to treat infections by S. aureus.
XX	
SQ	Sequence 9 AA;
XX	
Query Match	87.2%; Score 34; DB 21; Length 9;
Best Local Similarity	88.9%; Pred.No. 9.3e+05;
Matches	8; Conservative 1; Mismatches 0; Indels 0; Gaps 0.
OY	1 GVAXSALP 9 :
Db	1 GVAXSSLP 9
RESULT 12	
ID	AA67861 standard; peptide; 9 AA.
AC	AAV67861;
XX	
D7	25-APR-2000 ((first entry))
XX	
DE	Staphylococcus aureus AgrdII derived peptide sequence.
XX	
KW	Staphylococcus aureus infection; Agrd; agr response; treatment;
RW	virulence factor.
XX	
OS	Staphylococcus aureus.
XX	
FH	Key
FT	Misc-difference 5 Location/Qualifiers
PX	/label= Unknown
PN	
XX	
MO9967286-A2.	
PD	
29-DEC-1999.	
PF	24-JUN-1999; 99WO-US14562.

```

XX PR 24-JUN-1998; 98US-0103438.
XX XX (UYRO ) UNIV ROCKEFELLER.
XX PA (UNYV ) UNIV NEW YORK STATE.
XX PI Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
XX DR WPI; 2000-147202/13.
XX PT New cyclic peptides for treating infections with Staphylococcus aureus
XX PT
XX PS Examples; Page 22; 37pp; English.
XX CC This sequence represents the Staphylococcus aureus AgrdII derived
XX CC peptide. The invention relates to Agrd derived cyclic peptides; a
XX CC composition containing a peptide and a carrier, and a method for the
XX CC production of the cyclic peptides. The peptide inhibits the agr response,
XX CC which is normally associated with the release of virulence factors of
XX CC Staphylococcus aureus. An Agrd peptide is produced by S. aureus that
XX CC activates the agr response in strains of a single group, but interferes
XX CC with this response in strains of different groups. The peptides and
XX CC composition containing them can be used to treat infections by S. aureus.
SQ Sequence 9 AA;
Query Match 87.2%; Score 34; DB 21; Length 9;
Best Local Similarity 88.9%; Pred. No. 9.3e+05;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GVNAXSALF 9
DB 1 GVNAXSSLP 9
RESULT 13
ABP53540
ID ABP53540 standard; peptide; 9 AA.
XX AC ABP53540;
XX DT 13-DEC-2002 (first entry)
XX DE Cyclic peptide SEQ ID NO:1.
XX KW Cyclic peptide; Staphylococcus aureus; infection; antibacterial;
XX KM agr response inhibitor.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 5 /note= "any amino acid"
XX PN US2002077453-A1.
XX PD 20-JUN-2002.
XX PF 27-DEC-2001; 2001US-0032950.
XX PR 24-JUN-1998; 98US-090402P.
XX PR 24-JUN-1999; 99US-0339511.
XX PA (MUIR/) MUIR T W.
XX PA (MAYV/) MAYVILLE P.
XX PA (NOVI/) NOVICK R P.
XX PA (BEAV/) BEAVIS R.
XX PA (JIGG/) JI G.
XX PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX DR WPI; 2002-681366/73.

```

```

XX XX New cyclic peptides, useful for treating Staphylococcus aureus
XX PT infections
XX PS Claim 9; Page 10; 18pp; English.
XX CC ABP53540 to ABP53547 represent cyclic peptides (I) from the present
XX CC invention. The present invention also describes a method for treating
XX CC Staphylococcus aureus infection comprising the administration of a
XX CC composition comprising (I). (I) has antibacterial activity, and can be
XX CC used as an agr gene response inhibitor. The peptides are useful for
XX CC treating S. aureus infections.
SQ Sequence 9 AA;
Query Match 87.2%; Score 34; DB 23; Length 9;
Best Local Similarity 88.9%; Pred. No. 9.3e+05;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GVNAXSALF 9
DB 1 GVNAXSSLP 9
RESULT 14
AAM50899
ID AAM50899 standard; Peptide; 9 AA.
XX AC AAM50899;
XX DT 08-MAY-2002 (first entry)
XX DE Agrd-autoinducing cyclic peptide, inhibitor of agr response.
XX KM Staphylococcus aureus; Agrd; agr response; inhibitor; antibiotic;
XX KM antibacterial; infection; therapy; cyclic.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 5 /note= "any amino acid"
XX FT Misc-difference 5 /note= "note linked to residue 9 to form cyclic
XX FT peptide"
XX FT Misc-difference 9 /note= "note linked to residue 5 to form cyclic
XX FT peptide"
XX PN US6337385-B1.
XX PD 08-JAN-2002.
XX PF 24-JUN-1999; 99US-0339511.
XX PR 24-JUN-1998; 98US-090402P.
XX PA (UYRO ) UNIV ROCKEFELLER.
XX PA (UNYV ) UNIV NEW YORK STATE.
XX PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX DR WPI; 2002-170774/22.
XX PT Novel synthetic, cyclic Agrd-autoinducing peptide for bacterial
XX PT interference and for treating Staphylococcus aureus infection in a
XX PT subject
XX PS Claim 7; Column 19; 18pp; English.
XX CC The present sequence is that of a novel synthetic cyclic peptide
XX CC of the invention that is capable of inhibiting the agr response of
XX CC Staphylococcus aureus. It is an Agrd-autoinducing peptide, where

```


CC AgrD is a secreted agr-encoded peptide and where the agr locus
 CC controls the synthesis of virulence factor and other extracellular
 CC proteins responsible for pathogenicity in *S. aureus*. Preferred
 CC peptides may have the sequence NH₂-X(n)-Z-X(Y)-COOH, with a
 CC cyclic bond between the Z residue and COOH other than a thioester
 CC bond, where X is an amino acid, an amino acid analogue, a
 CC peptidomimetic or non-amide isostere, Z is a synthetic or a
 CC biosynthetic amino acid, n is 0-10 and Y is 1-10. The cyclic bond
 CC is especially a lactam or lactone bond. The thiololactone
 CC structure within native Agr peptides is required for activation
 CC of the agr response. Elimination of the thiol ester component of
 CC the cyclic ring structure can destroy agr response activating
 CC activity while preserving and enhancing inhibitory activity. A
 CC claimed method of preparing a cyclic peptide involves: assembling
 CC a linear peptide chain on to a solid phase resin support;
 CC deprotecting the resulting protected assembled peptide; treating the
 CC deprotected peptide with neutral buffer for a time sufficient to
 CC form the cyclic peptide and cleave the peptide from the support;
 CC and recovering the cyclic peptide. The peptide is useful for
 CC bacterial interference, especially for the treatment of *S. aureus*
 CC infection.

CC Sequence 9 AA;

Query Match 87.2%; Score 34; DB 23; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05; Indels 0; Gaps 0;
 Matches 8; Conservative 1; Mismatches 0;

QY 1 GVNXASALF 9
 |||||:
 Db 1 GVNXSSLF 9

RESULT 15
 AAMS0908
 ID AAMS0908 standard; Peptide; 9 AA.

AC AAMS0908;

DT 08-MAY-2002 (first entry)

DE Protected peptide used in cyclic peptide production.

KW *Staphylococcus aureus*; AgrD; agr response; inhibitor; antibiotic;
 KW antibacterial; infection; therapy; cyclic.

OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "Z-Gly"

FT Modified-site 5 /label= Dpr(Boc)

FT Modified-site 6 /note= "Ser(Bz1)"

FT Modified-site 7 /note= "Ser(Bz1)"

FT Misc-difference 5 /note= "note linked to residue 9 to form cyclic peptide"

FT Misc-difference 9 /note= "note linked to residue 5 to form cyclic peptide"

FT US6337385-B1.

XX 08-JAN-2002.

XX 24-JUN-1999; 99US-0339511.

XX 24-JUN-1998; 98US-090402P.

XX (UVRQ) UNIV ROCKEFELLER.

PA (UVRQ) UNIV NEW YORK STATE.
 XX Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
 PI MPI; 2002-170774/22.

PT Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
 PT interference and for treating *Staphylococcus aureus* infection in a
 PT subject -
 PS Disclosure; Column 14; 18pp; English.

CC The present sequence is that of a protected peptide used in an
 CC example of the preparation of novel synthetic cyclic peptides of
 CC the invention (see AAMS0899-906). The peptide corresponds to the
 CC *Staphylococcus aureus* AgrDII sequence with a Cys to diamino propionic
 CC acid (Dpr) mutation (lactam). It was synthesized on a Wang-resin
 CC using an Fmoc-N-alpha protection strategy. Following chain assembly,
 CC the peptide was cleaved from the support and the Dpr-5 residue
 CC deprotected by treatment with a trifluoroacetic acid:anisole:water
 CC mixture (50:5:5) for 4 hr. The partially protected peptide-alpha
 CC carboxylates were then dissolved in DMF and treated with PyBOP.
 CC Cyclization was complete after 2 hr. The remaining protecting groups
 CC were removed by treatment with HF and the peptide purified by HPLC.
 CC The cyclic peptide is capable of inhibiting the agr response of
 CC *Staphylococcus aureus*. The thiololactone structure within native
 CC AgrD peptides is required for activation of this response.
 CC Replacement of the thiol ester component of the cyclic ring
 CC structure with a lactam (as in the present case) or a lactone can
 CC destroy agr response activating activity while preserving and
 CC enhancing inhibitory activity. The cyclic peptides are useful for
 CC bacterial interference, especially for the treatment of *S. aureus*
 CC infection.

CC Sequence 9 AA;

Query Match 87.2%; Score 34; DB 23; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05; Indels 0; Gaps 0;
 Matches 8; Conservative 1; Mismatches 0;

QY 1 GVNXASALF 9
 |||||:
 Db 1 GVNXSSLF 9

Search completed: January 29, 2004, 14:52:13
 Job time : 36.0735 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:46:49 ; Search time 11.1176 Seconds
(without alignments)
77.851 Million cell updates/sec

Title: us-10-032-950-6

Perfect score: 39

Sequence: 1 GVNAXSALP 9

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR76:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	33	84.6	47	2	C89995	AgRP protein [limp
2	32	82.1	382	2	A28067	lysosomal membrane
3	32	82.1	405	2	A60534	P2B/LAMP-1 precurs
4	32	82.1	407	2	A30200	120K lysosomal mem
5	31	79.5	333	1	C69812	ferrichrome ABC tr
6	31	79.5	442	2	AF2539	manganese transport
7	31	79.5	1049	2	T30525	alpha-mannosidase
8	30	76.9	234	2	S60885	ferric exochelin u
9	30	76.9	298	2	T22250	hypothetical prote
10	30	76.9	361	2	AD3198	hypothetical prote
11	30	76.9	533	2	AG2328	2,3-bisphosphoglyc
12	30	76.9	549	2	P64640	conserved hypotetr
13	29	74.4	188	2	G84041	hypothetical prote
14	29	74.4	250	2	AB0421	Phnp protein [limp
15	29	74.4	255	2	AD1159	flagellar biosynth
16	29	74.4	255	2	AD1518	flagellar biosynth
17	29	74.4	259	2	S59439	transcription fact
18	29	74.4	346	2	S59491	probable galactosy
19	29	74.4	605	2	H71562	probable flagellar
20	29	74.4	895	2	T45786	receptor-protein k
21	29	74.4	1016	2	T30942	aminopeptidase (EC
22	29	74.4	1016	2	T30943	aminopeptidase (EC
23	29	74.4	1421	2	T34325	hypothetical prote
24	28	71.8	43	2	S20040	hypothetical prote
25	28	71.8	43	2	S20039	hypothetical prote
26	28	71.8	148	2	D97308	probable membrane
27	28	71.8	167	2	G87610	hypothetical prote
28	28	71.8	259	2	H75481	hydroxylase, CbbY/Cb
29	28	71.8	303	2	D84230	hypothetical prote

30	28	71.8	341	2	C89996	hypothetical prote
31	28	71.8	346	2	F69786	glycoprotein endop
32	28	71.8	389	1	S36653	kpsd protein - Bsc
33	28	71.8	394	2	E81286	probable polysacch
34	28	71.8	452	2	F64053	4-chlorobenzate-C
35	28	71.8	468	1	WMBE42	DNA-binding protei
36	28	71.8	491	2	D95365	NADH2 dehydrogenas
37	28	71.8	546	2	A32260	cholesterol oxidas
38	28	71.8	573	2	F81313	peptidase (Mj fami
39	28	71.8	631	2	D69750	phosphotransferase
40	28	71.8	655	2	G89189	protein y32F6a.3 (
41	28	71.8	803	2	T39530	hypothetical prote
42	28	71.8	832	2	S46706	probable transcrip
43	28	71.8	870	2	S74291	hypothetical prote
44	28	71.8	926	2	S49463	SEC24 protein - Ye
45	28	71.8	1844	1	RRWFTW	genome polyprotein

ALIGNMENTS

RESULT 1
C89995
AgRP protein [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89995
R:Kurita, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oe, M.; A.; Mizutani-Uh, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.; C.; Shiba, T.; Hatton, M.; Ogasawara, N.; Hayashi, H.; Hirataesu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: C89995
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-47 <NR>
A:Cross-references: GB:BA000018; PID:G13701831; PIDN:BA83124.1; GSPDB:GNO0149
A:Experimental source: strain N315
C:Genetics:
A:Gene: agrD

Query Match 84.6%; Score 33; DB 2; Length 47;
Best Local Similarity 77.8%; Pred. No. 0.87;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALP 9
DB 24 GVNAXSALP 32

RESULT 2
A28067
lysosomal membrane glycoprotein LAMP-1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 20-Aug-1999
C:Accession: A28067
R:Chen, J.W.; Cha, Y.; Yuxsel, K.U.; Gracy, R.W.; August, J.T.
J. Biol. Chem. 263, 8754-8758, 1988
A:Title: Isolation and sequencing of a cDNA clone encoding lysosomal membrane glycopr
A:Reference number: A28067; PMID:88243732; PMID:3379044
A:Accession: A28067
A:Molecule type: mRNA
A:Residues: 1-382 <CH>
A:Cross-references: GB:J03881; NID:G198706; PIDN:AAA39411.1; PID:G293682
A:Note: the authors translated the codon ATT for residue 1 as Leu and CCG for residue
C:Superfamily: lysosome-associated membrane protein
C:Keywords: glycoprotein; membrane protein

Query Match 82.1%; Score 32; DB 2; Length 382;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
Db 256 GVNASSSLF 264

RESULT 3

A60534
P2B/LAMP-1 precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 03-May-1996
C/Accession: A60534
R/Hefemann, M.; Yousefi, S.; Dennis, J.W.
Cancer Res. 49, 6077-6084, 1989
A/Title: Molecular characterization of P2B/LAMP-1, a major protein target of a metastasizing melanoma cell line
A/Accession: A60534; MUID:90002989; PMID:2676155
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-405 <HEF>
C/Superfamily: lysosome-associated membrane protein

Query Match 82.1%; Score 32; DB 2; Length 405;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
Db 279 GVNASSSLF 287

RESULT 4

A30200
120K lysosomal membrane glycoprotein precursor - rat
M/Alternate names: sialoglycoprotein
C/Species: Rattus norvegicus (Norway rat)
C/Date: 18-Oct-1989 #sequence_revision 30-Sep-1991 #text_change 20-Aug-1999
C/Accession: A30200; S03331
R/Howe, C.L.; Granger, B.L.; Hull, M.; Green, S.A.; Gabel, C.A.; Helenius, A.; Mellan, Proc. Natl. Acad. Sci. U.S.A. 85, 7577-7581, 1988
A/Title: Derived protein sequence, oligosaccharides, and membrane insertion of the 120-K glycoproteins.
A/Reference number: A30200; MUID:89017240; PMID:3174552
A/Accession: A30200
A/Molecule type: mRNA
A/Residues: 1-407 <HOW>
A/Cross-references: EMBL:J03672
A/Note: the authors translated the codon GGG for residue 15 as Val
R/Henno, M.; Noguchi, Y.; Sasaki, H.; Tanaka, Y.; Furuno, K.; Kono, A.; Sakaki, Y.; Kat FEMS Lett. 244, 351-356, 1989
A/Title: Isolation and sequencing of a cDNA clone encoding 107 kDa sialoglycoprotein in A/Reference number: S03331; MUID:89153580; PMID:2920835
A/Accession: S03331
A/Molecule type: mRNA
A/Residues: 22-407 <HIM>
A/Cross-references: EMBL:X14765; NID:G56577; PIDN:CA32873.1; PID:G56578
A/Note: part of this sequence, including the amino end of the mature protein, was confirmed by the authors
C/Superfamily: lysosome-associated membrane protein
C/Keywords: glycoprotein; membrane protein
F/22-407/Product: lysosomal membrane glycoprotein, 107K #status experimental <MAT>

Query Match 82.1%; Score 32; DB 2; Length 407;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
Db 281 GVNATSSLF 289

RESULT 5
C69812
ferrichrome ABC transporter (permease) homolog yfmd - Bacillus subtilis
C/Species: Bacillus subtilis

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C/Accession: C69812

R/Kunze, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berl C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; C.A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall, A.; Harwood, C.R.; Hensley, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M. Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau Y. M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portee, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanli A.; Authors: Schleier, S.; Schroeder, R.; Scrofano, F.; Sekiguchi, J.; Sekowska, A.; Set akuch, M.; Tanakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiy T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, A.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: C69812
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-333 <KUN>
A/Cross-references: GB:Z99108; GB:AL009126; NID:G2633055; PIDN:CA312580.1; PID:G26330
A/Experimental source: strain 168
C/Genetics:
A/Gene: yfmd
C/Superfamily: ferrichrome ABC transporter

Query Match 79.5%; Score 31; DB 1; Length 333;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
Db 99 GVNASSSLF 107

RESULT 6

AF2539
manganese transport protein all7601 [imported] - Nostoc sp. (strain PCC 7120) plasmid
C/Species: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C/Accession: AF2539
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriugu Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata DNA Res. 8, 205-213, 2001
A/Title: Complete genomic sequence of the filamentous nitrogen-fixing Cyanobacterium A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: AF2539
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-442 <KUR>
A/Cross-references: GB:AP003602; PIDN:BA87244.1; PID:G1734686; GSPDB:GN00181
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: all7601
C/Superfamily: natural resistance-associated macrophage protein 1

Query Match 79.5%; Score 31; DB 2; Length 442;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
Db 301 GVNASSSLF 309

RESULT 7
T30525
alpha-mannosidase (EC 3.2.1.24) - Emericella nidulans
C/Species: Emericella nidulans, Aspergillus nidulans

C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
 C/Accession: T30525
 R/Pades, C.J.; Gilbert, A.; Goodman, C.D.; Hintz, W.E.
 C/Species: *Caenorhabditis elegans*
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T22250
 R/Kelly, P.
 submitted to the EMBL Data Library, August 1996
 A/Reference number: Z19537
 A/Accession: T22250
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-298 <WIL>
 A/Cross-references: EMBL:Z78420; PIDN:CAB01709.1; GSPDB:GN00019; CESP:F45H11.1
 A/Experimental source: clone F45H11
 C/Genetics:
 A/Introns: 49/2; 219/1; 877/2
 A/Note: msd2
 C/Superfamily: Saccharomyces alpha-mannosidase
 C/Keywords: glycosidase; hydrolase

Query Match 79.5%; Score 31; DB 2; Length 1049;
 Best Local Similarity 87.5%; Pred. No. 62;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVMAXSALF 9
 Db 979 GVAVSALF 966

RESULT 8
 S60885
 ferric exochelin uptake protein fxuC - *Mycobacterium smegmatis* (fragment)
 C/Species: *Mycobacterium smegmatis*
 C/Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 22-Oct-1999
 C/Accession: S60885
 R/Pfister, E.H.; Yu, S.; Jacobs Jr., W.R.
 M.O.I. Microbiol. 14, 557-569, 1994
 A/Title: Identification of genes involved in the sequestration of iron in mycobacteria:
 A/Reference number: S60885; MUID:95191405; PMID:7885234
 A/Accession: S60885
 A/Molecule type: DNA
 A/Residues: 1-234 <FIS>
 A/Cross-references: EMBL:U10425; NID:G595400; PIDN:AAQ43258.1; PID:G595401
 C/Genetics:
 A/Gene: fxuC

Query Match 76.9%; Score 30; DB 2; Length 234;
 Best Local Similarity 55.6%; Pred. No. 23;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVMAXSALF 9
 Db 37 GINSALF 45

RESULT 9
 T22250
 hypothetical protein F45H11.1 - *Caenorhabditis elegans*
 C/Species: *Caenorhabditis elegans*
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T22250
 R/Kelly, P.
 submitted to the EMBL Data Library, August 1996
 A/Reference number: Z19537
 A/Accession: T22250
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-298 <WIL>
 A/Cross-references: EMBL:Z78420; PIDN:CAB01709.1; GSPDB:GN00019; CESP:F45H11.1
 A/Experimental source: clone F45H11
 C/Genetics:
 A/Introns: 52/3; 109/2; 177/3; 248/1; 290/3

Query Match 76.9%; Score 30; DB 2; Length 298;

Best Local Similarity 75.0%; Pred. No. 29;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVMAXSAL 8
 Db 159 GINAGSAL 166

RESULT 10
 AD3198
 hypothetical protein fepd [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)
 C/Species: *Agrobacterium tumefaciens*
 C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C/Accession: AD3198
 R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCle ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm ster, E.W.
 A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
 A/Reference number: AB2577; MUID:21608550; PMID:11743193
 A/Accession: AD3198
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-361 <KUB>
 A/Cross-references: GB:AB008687; PIDN:AA46002.1; PID:G17743757; GSPDB:GN00188
 A/Experimental source: strain C58 (Dupont)
 C/Genetics:
 A/Gene: fepd
 A/Genome: plasmid

Query Match 76.9%; Score 30; DB 2; Length 361;
 Best Local Similarity 66.7%; Pred. No. 36;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVMAXSALF 9
 Db 122 GVNAGAAFF 130

RESULT 11
 AG2328
 2,3-bisphosphoglycerate-independent phosphoglycerate mutase [imported] - *Nostoc* sp. (i
 C/Species: *Nostoc* sp. PCC 7120
 A/Note: Nostoc sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
 C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C/Accession: AG2328
 R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguc Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata. DNA Res. 8, 205-213, 2001
 A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing *Cyanobacterium* ;
 A/Reference number: AB1807; MUID:21595285; PMID:11759840
 A/Accession: AG2328
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-533 <KUB>
 A/Cross-references: GB:BA000019; PIDN:BA75881.1; PID:G17133317; GSPDB:GN00179
 A/Experimental source: strain PCC 7120
 C/Genetics:
 A/Gene: al1482
 C/Superfamily: phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent

Query Match 76.9%; Score 30; DB 2; Length 533;
 Best Local Similarity 75.0%; Pred. No. 53;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVMAXSAL 8
 Db 163 GINAGSAL 170

RESULT 12

F64640 conserved hypothetical protein HP0966 - Helicobacter pylori (strain 26695)
 C:Species: Helicobacter pylori
 C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
 C:Accession: F64640
 R:Tombl, J.F.; White, O.; Kervarage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khatkhat, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
 A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A:Reference number: A64520; MUID:97394467; PMID:9252185
 A:Accession: F64640
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-549 <TOM>
 A:Cross-references: GB:AE000605; GB:AE000511; NID:G2314103; PIDN:AAD08010.1; PID:G231410

Query Match 76.4%; Score 30; DB 2; Length 549;
 Best Local Similarity 55.6%; Pred. No. 55;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
 |||:|:
 Db 406 GINLSVTF 414

RESULT 13

GB4041 hypothetical protein BH3135 [imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: GB4041
 R:Takami, H.; Nakase, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Maeni, N.; Fuji, F.; Hira, Nucleic Acids Res. 28, 4317-4331, 2000
 A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A63650; MUID:20512582; PMID:11058132
 A:Accession: GB4041
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-188 <STC>
 A:Cross-references: GB:AP001517; GB:BA000004; NID:G10175500; PIDN:BA006854.1; GSPDB:GN00
 A:Experimental source: strain C-125
 C:Genetics:

Query Match 74.4%; Score 29; DB 2; Length 188;
 Best Local Similarity 66.7%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
 |||:|:
 Db 60 GLTATSAF 68

RESULT 14

AB0421 PhnP protein [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Dec-2002
 C:Accession: AB0421
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Ill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett, Nature 413, 523-527, 2001
 A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AB0421
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-250 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC92694.1; PID:G15981389; GSPDB:GN00175
 C:Genetics:

A:Gene: phnP
 C:Superfamily: ATP-binding protein PhnP (PhnP)

Query Match 74.4%; Score 29; DB 2; Length 250;
 Best Local Similarity 66.7%; Pred. No. 42;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
 |||:|:
 Db 165 GLSADSAF 173

RESULT 15

AD1159 flagellar biosynthetic protein flp homolog lmo676 [imported] - Listeria monocytogenes
 C:Species: Listeria monocytogenes
 C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 C:Accession: AD1159
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fehl, D.; Jones, L.M.; Karsch, U. Science 294, 849-852, 2001
 A:Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maitournam, A.; Ok, C.; Schluter, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla, A>Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AD1159
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-255 <GLA>
 A:Cross-references: GB:NC_003210; PIDN:CAC98754.1; PID:G16410065; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:

A:Gene: lmo676
 C:Superfamily: flagellar biosynthetic protein flp

Query Match 74.4%; Score 29; DB 2; Length 255;
 Best Local Similarity 66.7%; Pred. No. 43;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
 |||:|:
 Db 47 GVNSSVALF 55

Search completed: January 29, 2004, 14:58:23
 Job time : 12.1176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:45:44 ; Search time 5.82353 Seconds

(without alignments)
72.678 Million cell updates/sec

Title: US-10-032-950-6

Perfect score: 39

Sequence: 1 GUNMAXSAUF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Length	ID	Description
1	82.1	406	1 LMP1_MOUSE	P11438 mus muscula
2	82.1	407	1 LMP1_RAT	P14562 ratu mus norv
3	79.5	442	1 MNTH_ANASP	Q84560 anabena sp
4	76.9	533	1 GPM1_ANASP	O8YPI2 anabena sp
5	74.4	269	1 AKR_CHICK	O90655 gallus gall
6	74.4	346	1 YARD_SCHPO	O09681 schistosach
7	71.8	117	1 RBPA_LEPIN	O68770 leptospira
8	71.8	243	1 RECO_VIBVU	Q08674 vibrio vuln
9	71.8	303	1 PANE_HAHLN1	O09010 halobacteri
10	71.8	346	1 GCP_BACSV	O05518 bacillus su
11	71.8	389	1 KSS5_ECOLI	P44218 escherichia
12	71.8	407	1 LMP1_CRIGR	P49129 cricetus
13	71.8	452	1 MENE_HAEIN	P44565 haemophilus
14	71.8	488	1 VPAP_HSV11	P10222 herpes simp
15	71.8	546	1 CHOD_STRSQ	P12676 streptomyc
16	71.8	631	1 YBFS_BACSV	P33816 bacillus su
17	71.8	870	1 YCSO_YEAST	P25623 saccharomyc
18	71.8	883	1 YH6_YEAST	P38781 saccharomyc
19	71.8	926	1 SC24_YEAST	P40482 saccharomyc
20	71.8	1844	1 POLR_TYMV	P13358 turnip yell
21	71.8	1844	1 POLR_TYMV	P20128 turnip yell
22	71.8	1844	1 POLR_TYMV	P20477 turnip yell
23	71.8	1844	1 POLR_TYMV	P20477 turnip yell
24	71.8	1844	1 POLR_TYMV	P20477 turnip yell
25	71.8	1844	1 POLR_TYMV	P20477 turnip yell
26	71.8	1844	1 POLR_TYMV	P20477 turnip yell
27	71.8	1844	1 POLR_TYMV	P20477 turnip yell
28	71.8	1844	1 POLR_TYMV	P20477 turnip yell
29	71.8	1844	1 POLR_TYMV	P20477 turnip yell
30	71.8	1844	1 POLR_TYMV	P20477 turnip yell
31	71.8	1844	1 POLR_TYMV	P20477 turnip yell
32	71.8	1844	1 POLR_TYMV	P20477 turnip yell
33	71.8	1844	1 POLR_TYMV	P20477 turnip yell

ALIGNMENTS

RESULT 1
ID LMP1_MOUSE STANDARD; PRT; 406 AA.
AC P11438; 062020; (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 15, Last sequence update)
DT 01-AUG-1990 (Rel. 35, Last annotation update)
DE Lysosome-associated membrane glycoprotein 1 precursor (LAMP-1) (LGP-A)
DE (LGP-120) (CD107A) (P2B).
GN LAMP1 OR LAMP-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90307738; PubMed=2142158;
RA Granger B.L., Green S.A., Gabel C.A., Howe C.L., Mellman I.,
RA Helenius A.;
RT "Characterization and cloning of 19p110, a lysosomal membrane
RT glycoprotein from mouse and rat cells.";
RL J. Biol. Chem. 265:12036-12043 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RA Hofferman M., Yousefi S., Dennis J.W.;
RL Submitted (FEB-1990) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 25-406 FROM N.A.; AND PARTIAL SEQUENCE.
RX MEDLINE=88243732; PubMed=3379044;
RA Chen J.W., Cha Y., Yukeal K.U., Gracy R.W., August J.T.;
RT "Isolation and sequencing of a cDNA clone encoding lysosomal membrane
RT glycoprotein mouse LAMP-1. Sequence similarity to proteins bearing
RT onco-differentiation antigens.";
RL J. Biol. Chem. 263:8754-8758 (1988).
RN [4]
RP DISULFIDE BONDS.
RX MEDLINE=90237040; PubMed=2332434;
RA Ardenburn L.W., Farles B.J., August J.T.;
RT "The disulfide structure of mouse lysosome-associated membrane
RT protein 1.";
RL J. Biol. Chem. 265:7419-7423 (1990).
CC -1- FUNCTION: PRESENTS CARBOHYDRATE LIGANDS TO SELECTINS. ALSO
CC IMPLICATED IN TUMOR CELL METASTASIS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Lysosomal.
CC THIS PROTEIN SHUTTLES BETWEEN LYSOSOMES, ENDOSOMES, AND THE PLASMA
CC MEMBRANE.
CC -1- PTM: O- AND N-GLYCOSYLATED, SOME OF THE N-GLYCANS ATTACHED TO
CC LAMP-1 ARE POLYLACTOSAMINOGLYCANS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LAMP FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)


```

FT CAROHD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 223 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 293 294 GV -> EF (IN REF. 3).
FT CONFLICT 329 329 N -> T (IN REF. 3).
FT CONFLICT 356 357 SD -> VT (IN REF. 3).
SQ SEQUENCE 407 AA; 43969 MW; 25947490749A7C98 CRC64;

Query Match
Best Local Similarity 82.1%; Score 32; DB 1; Length 407;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
DB 281 GVNATSLF 289

RESULT 3
MNTN ANASP STANDARD; PRT; 442 AA.
AC 0823B0:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable manganese transport protein mntn.
GN MNTN OR AL17601.
OS Anabaena sp. (strain PCC 7120).
OC Plasmid pCC7120beta.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneo T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Matenabe A., Iriuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
"Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
DNA Res. 8:205-213(2001).
RT FUNCTION: H(+) -stimulated, highly selective, manganese uptake
system (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(CC) (Probable).
CC -1- SIMILARITY: BELONGS TO THE NRAMP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP003602; BAB77244.1; -
DR PIR; AP2539; AF2539.
DR HAMAP; MF_00221; -; 1.
DR InterPro; IPR001046; Nramp.
DR Pfam; PF01566; Nramp; 1.
DR PRINTS; PR00447; NAIRESASSCMP.
DR PRODOM; PD001861; Nramp; 1.
DR TIGRFAWS; TIGR01197; nramp; 1.
KW Transport; Symport; Manganese; Transmembrane; Inner membrane; Plasmid;
KW Complete proteome.
FT TRANSMEM 29 49 POTENTIAL.
FT TRANSMEM 62 82 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.

```

```

FT TRANSMEM 135 155 POTENTIAL.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 209 229 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
FT TRANSMEM 295 315 POTENTIAL.
FT TRANSMEM 347 367 POTENTIAL.
FT TRANSMEM 372 392 POTENTIAL.
FT TRANSMEM 413 433 POTENTIAL.
SQ SEQUENCE 442 AA; 48418 MW; A7CBADFAE628196B CRC64;

Query Match
Best Local Similarity 79.5%; Score 31; DB 1; Length 442;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
DB 301 GVNAXSALF 309

RESULT 4
GMNT ANASP STANDARD; PRT; 533 AA.
ID GMNT ANASP
AC QBYPL2:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 2,3-bisphosphoglycerate-independent phosphoglycerate mutase
DE (EC 5.4.2.1) (Phosphoglyceromutase) (BPG-independent PGM) (PGM).
GN GPMI OR AL14182.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneo T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Matenabe A., Iriuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
"Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
DNA Res. 8:205-213(2001).
RT FUNCTION: Catalyzes the interconversion of 2-phosphoglycerate and
3-phosphoglycerate (By similarity).
CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = 3-phospho-D-glycerate;
CC -1- COFACTOR: Binds 2 manganese ions (By similarity).
CC -1- PATHWAY: Glycolysis.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SIMILARITY: BELONGS TO THE BPG-INDEPENDENT PHOSPHOGLYCERATE MUTASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP003595; BAB75881.1; -
DR PIR; AG2328; AG2328.
DR HAMAP; MF_01038; -; 1.
DR InterPro; IPR006124; Metalloenzyme.
DR InterPro; IPR005995; Pgm_bpd_ind.
DR Pfam; PF01676; Metalloenzyme; 1.
DR PRODOM; PD004429; Pgm_bpd_ind; 1.
DR TIGRFAWS; TIGR01307; Pgm_bpd_ind; 1.
KW Phosphatase; Glycolysis; Metal-binding; Manganese; Complete proteome.
KW ACT_SITE 65 65 PHOSPHOGLYCERATE INTERMEDIATE
FT METAL 15 15 (BY SIMILARITY).
FT METAL 65 65 MANGANESE 2 (BY SIMILARITY).

```


FT METAL 398 398 MANGANESE 1 (BY SIMILARITY).
 FT METAL 402 402 MANGANESE 1 (BY SIMILARITY).
 FT METAL 439 439 MANGANESE 2 (BY SIMILARITY).
 FT METAL 440 440 MANGANESE 2 (BY SIMILARITY).
 FT METAL 457 457 MANGANESE 1 (BY SIMILARITY).
 SQ SEQUENCE 533 AA; 57677 MW; ACC40C19F612A97C CRC64;

Query Match 76.9%; Score 30; DB 1; Length 533;
 Best Local Similarity 75.0%; Pred. No. 26;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GMAXSALF 8
 DB 163 GMAXSALF 170

RESULT 5

AKR_CHICK STANDARD; PRT; 269 AA.
 ID AKR_CHICK
 AC Q090551
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeodomain protein AKR (avian knotted-related protein).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RX MEDLINE=95396587; PubMed=7667102;
 RP TISSUE=Liver;
 RC MEDLINE=95396587; PubMed=7667102;
 RA Ryan A.K., Tejeda M.L., May D.L., Dubrava M., Delsey R.G.;
 RT "Isolation and characterization of the chicken homeodomain protein AKR";
 RT Nucleic Acids Res. 23:3252-3259 (1995).
 CC -1- FUNCTION: BINDS TO THE P' ELEMENT OF THE APOVDL1 GENE AND REPRESS ITS TRANSCRIPTION.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE TRUE/TGIF HOMEBOX FAMILY.
 CC -1- SIMILARITY: Contains 1 homeobox domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; U25353; AAA83567.1; -
 DR PIR; S58439; S58439.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; FALSE_NEG.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 KW Transcription regulation; Repressor; Homeobox; DNA-binding; Nuclear protein.
 FT DNA BIND 35
 FT SEQUENCE 269 AA; 29442 MW; 1074355DCC0C2253 CRC64;

Query Match 74.4%; Score 29; DB 1; Length 269;
 Best Local Similarity 66.7%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GMAXSALF 9
 DB 222 GMAXSALF 230

RESULT 6
 YA0D SCHPO STANDARD; PRT; 346 AA.
 ID YA0D SCHPO
 AC C09681;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative glycyl transferase C5H10.13c in chromosome I (BC 2.-.-.-).
 GN SPAC5H10.13c.
 OS Schizosaccharomyces pombe (Fission Yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OC NCBI_TaxID=4896;
 RX MEDLINE=21648401; PubMed=11859360;
 RC STRAIN=972;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagers K., James K., Jones S., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volkhardt G., Aert R., Robben J., Grymoprez B., Wellens I., Vanstreels E., Rieger M., Scheffer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Leinhardt R., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Medler H., Wambutt R., Punnelle B., Goffeau A., Gadieu E., Deyano S., Gloux S., Leleau V., Motier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hunt S.M., Lucas R., Rochet M., Gallard J., Tallada V.A., Garzon A., Thode G., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Fetsburg S.L., Derrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Sipakovski G.V., Useery D., Barrett B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880 (2002).
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (potential).
 CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 34.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; Z49811; CAAG9963.1; -
 DR PIR; T38977; S55491.
 DR GeneDB; Spome; SPAC5H10.13c; -
 KW Hypothetical protein; Transferase; Glycosyltransferase; Transmembrane; Signal-anchor; Golgi stack; Glycoprotein.
 FT DOMAIN 1
 FT TRANSMEM 12
 FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).
 FT LUMENAL (POTENTIAL).
 FT CARBOHYD 33
 FT CARBOHYD 64
 FT CARBOHYD 142
 FT CARBOHYD 224
 SQ SEQUENCE 346 AA; 39107 MW; 1AD26CDDABAC0D3B CRC64;

Query Match 74.4%; Score 29; DB 1; Length 346;
 Best Local Similarity 66.7%; Pred. No. 28;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

QY      1 GVNAXSALF 9
DB      243 GUNASGLF 251

RESULT 7
RBPAL LEPIN STANDARD; PRT; 117 AA.
AC 08F7K0;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ribosome-binding factor A.
GN RBPAL OR RBPAL OR LAD344.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
CX NCBI_TaxID=173;

SEQUENCE FROM N.A.
RP SRRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA Zhang Y.-X., Xiong H., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
RA Zhang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA Saint Girons I., Somerville R.L., Wen Y.-W., Shi W.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT Unique physiological and pathogenic features of Leptospira
RT Interrogans revealed by whole-genome sequencing."
RL Nature 422:888-893 (2003).
CC -1- FUNCTION: Associates with free 30S ribosomal subunits (but not
CC with 30S subunits that are part of 70S ribosomes or polysomes).
CC Essential for efficient processing of 16S rRNA. May interact with
CC the 5' terminal helix region of 16S rRNA (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO THE RBPAL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE011278; AAN48143.1; -
DR HAMAP: MF 00003; -; 1.
DR Pfam: PF02033; RBPAL; 1.
DR ProDom: PD007327; Rib_bind_facA; 1.
DR TIGRFAMs: TIGR00082; TdFA; 1.
DR PROSITE: PS01319; RBPAL; 1.
KM rRNA processing; Complete proteome.
SQ SEQUENCE 117 AA; 13252 MW; D304576B560DEB89 CRC64;

Query Match 71.8%; Score 28; DB 1; Length 117;
Best Local Similarity 55.6%; Pred. No. 15;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
CX NCBI_TaxID=672;

SEQUENCE FROM N.A.
RP STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6."
RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Involved in DNA repair and recF pathway recombination
CC (By similarity).
CC -1- SIMILARITY: BELONGS TO THE RECO FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE016802; AAC09931.1; -
DR HAMAP: MF 00201; -; 1.
DR InterPro: IPR003717; RECO.
DR Pfam: PF02565; RECO; 1.
KM DNA repair; DNA recombination; Complete proteome.
SQ SEQUENCE 243 AA; 27570 MW; 65E5979525D1F8E CRC64;

Query Match 71.8%; Score 28; DB 1; Length 243;
Best Local Similarity 55.6%; Pred. No. 33;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 GVNAXSALF 9
DB      86 GINVSALY 94

RESULT 9
PANE HALNI STANDARD; PRT; 303 AA.
AC Q9H8F0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative 2-dehydropantoate 2-reductase (EC 1.1.1.169) (Ketopantoate
DE reductase) (KPR reductase) (KPR).
GN VNG0730C.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
CX NCBI_TaxID=64091;
RN [1]

SEQUENCE FROM N.A.
RP MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Maharis G.G., Bergquist B., Pan M.,
RA Shukla H.D., Laskey S.R., Balliga N.S., Thorsson V., Shrogha J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauer B., Keller K., Cruz R., Danon M.J., Hough D.W.,
RA Maddock D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Jung K.-H.,
RA Isenbarger T.A., Beck R.F., Pohlischer M., Spudich J.L., Dale K.-H.,
RA Alam W., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181 (2000).
CC -1- FUNCTION: Catalyzes the NADPH-dependent reduction of ketopantoate
CC into pantoic acid (By similarity).
CC -1- CATALYTIC ACTIVITY: (R)-pantoate + NADP(+) = 2-dehydropantoate +
CC NADPH.
CC -1- PATHWAY: Pantothenate biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO THE KETOPANTOATE REDUCTASE FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; A005017; AAC19208.1; -
CC PIR; D84230; D84230.
CC InterPro: IPR003710; Abpa.
CC Pfam; PF05558; Abpa.1.
CC TIGRfam; TIGR00745; abpa_pam.1.
CC Hypothetical protein; Pantothenate biosynthesis; Oxidoreductase; NADP;
CC Complete proteome.
CC NP BIND 7 12 NADP (POTENTIAL).
CC ACT_SITE 185 BY SIMILARITY.
CC ACT_SITE 267 BY SIMILARITY.
CC SEQUENCE 303 AA; 31124 MW; B373143f2822DC94 CRC64;
SQ
Query Match
Best Local Similarity 71.8%; Score 28; DB 1; Length 303;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
CY 1 GVMXSAL 8
DB 191 GINAXAL 198
RESULT 10
GCP_BACSU STANDARD; PRT; 346 AA.
AC 005518;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable O-sialoglycoprotein endoproteidase (EC 3.4.24.57)
DE (Glycoprotease).
GN GCP.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX NCBI_TaxID=1423;
RX STRAIN=168 / JH642;
RX MEDLINE=9346038; PubMed=9202461;
RX Sadate Y., Yata K., Fujita M., Sagai H., Itaya M., Kasahara Y.,
RX Ogasawara N.;
RT "Nucleotide sequence and analysis of the phoB-rnm-groESL region of
RT the Bacillus subtilis chromosome.";
RL Microbiology 143:1861-1866 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RC MEDLINE=9804033; PubMed=9384377;
RC Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RC Azevedo V., Bertorello M.G., Bessieres P., Bolotin A., Borchert S.,
RC Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RC Brouillet S., Bruschi C.V., Caldwell B., Capriano V., Carter N.M.,
RC Choi S.K., Codani U.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RC Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RC Enlian K.D., Eriington K.M., Fabret C., Ferrari E., Foulger D.,
RC Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RC Gilm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RC Giuseppe G., Guy B.J., Haga K., Hatach U., Harwood C.R., Henaut A.,
RC Hilbert H., Holappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RC Joris B., Karamata D., Kasahara Y., Klaerr-Bianhard M., Klein C.,
RC Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,
RC Kunita K., Lepidus A., Lardinois S., Lauber U., Lazarevic V.,
RC Lee S.M., Levine A., Liu H., Maeda S., Manuel C., Medigue C.,
RC Medina N., Meliadi R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RC Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

```

```

RA Pario V., Pohl T.M., Portetelle D., Portetelle S., Prescott A.M.,
RA Pressac E., Pyjic P., Purnelle B., Rappoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha S., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlon E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Serron S.J., Serron P., Shin B.S., Soldo B.,
RA Sorokin A., Taccani E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Tarpley P., Tognoni A.,
RA Toso V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters K., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zimstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis";
RL Nature 390:249-256 (1997).
CC -1- FUNCTION: COULD BE A METALLOPROTEASE.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of O-sialoglycoproteins; cleaves
CC 31-Arg-1-Asp-12 bond in glycoprotein A. Does not cleave
CC unglycosylated proteins, desialylated glycoproteins or
CC glycoproteins that are only N-glycosylated.
CC -1- CORFACTOR: ZINC (PROBABLE).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M22.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D88802; BAA19718.1; -
CC EMBL; Z91017; CAB12413.1; -
CC PIR; F69786; F69786.
CC MEROPS; M22; UPW; -
CC Subtilisin; BG12202; GCP.
CC InterPro: IPR009005; Peptidase_M22.
CC Pfam; PF00814; Peptidase_M22; 1.
CC PRINTS; PR00789; OSIALOPTASE.
CC Prodom; PD002167; Peptidase_M22; 1.
CC TIGRfam; TIGR00329; gcp.1.
CC PROSITE; PS01016; GLYCOPROTEASE; 1.
CC Hydrolase; Metalloprotease; Zinc; Complete proteome.
CC METAL 117 ZINC (POTENTIAL).
CC METAL 121 ZINC (POTENTIAL).
CC SEQUENCE 346 AA; 36842 MW; 0C7F618780DBA1A CRC64;
SQ
Query Match
Best Local Similarity 71.8%; Score 28; DB 1; Length 346;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CY 1 GVMXSAL 8
DB 97 GVMAXAL 104
RESULT 11
KSS5_ECOLI STANDARD; PRT; 389 AA.
AC K42218;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Capsule polysaccharide export protein Kps5.
GN KPS5.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
CX NCBI_TaxID=562;
RX STRAIN=K5;
RX MEDLINE=93388530; PubMed=8397187;
RX Pazani C., Rosenow C., Boulnois G.J., Bronner D., Jann K.,

```

```

RA Roberts.S.;
RT "Molecular analysis of region 1 of the Escherichia coli K5 antigen
RT gene cluster: a region encoding proteins involved in cell surface
RT expression of capsular polysaccharide."
RL J. Bacteriol. 175:1578-1583(1993).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74567; CAA52659.1;
DR PIR: S36653; S36653.
DR Pfam: PF05159; Capsule_synth; 1.
KW Polysaccharide transport; Transport.
SQ SOURCE 389 AA; 46381 MW; 73058122C28027DE CRC64;
QY
QY 1 GYMAXSAL 8
QY |||||
QY 127 GYVAYSSL 134
Db
Db 127 GYVAYSSL 134
RESULT 12
LMP1_CRIGR STANDARD; PRT; 407 AA.
AC
AC P49129;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Lysosomal-associated membrane glycoprotein 1 precursor (LAMP-1)
DE (lysosomal membrane glycoprotein A) (LGP A).
GN LAMP1 OR LGPA.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
OX 11
RN
RN SEQUENCE FROM N.A.
RP MEDLINE=97021428; PubMed=8667788;
RA Uthayakumar S., Granger B.L.;
RA "Cell surface accumulation of overexpressed membrane lysosomal
RT membrane glycoproteins";
RL Cell. Mol. Biol. Res. 41:405-420(1995).
CC -1- FUNCTION: PRESENTS CARBOHYDRATE LIGANDS TO SELECTINS. ALSO
CC IMPLICATED IN TUMOR CELL METASTASIS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Lysosomal.
CC THIS PROTEIN SHUTTLES BETWEEN LYOSOMES, ENDOSOMES, AND THE PLASMA
CC MEMBRANE (BY SIMILARITY).
CC -1- PTM: O- AND N-GLYCOSYLATED. SOME OF THE N-GLYCANS ATTACHED TO
CC LAMP-1 ARE POLYLACTOSAMINOGLYCANS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LAMP FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L18986; AAC37682.1;
DR InterPro: IPR002000; Lamp.
DR Pfam: PF01299; Lamp; 1.
DR PRINTS: PRO0036; LYSASSOCTDMP.
DR PROSITE: PS000310; LAMP 1; 2.

```

Qy	Db	Query Match	Best Local Similarity	Score	Length	Matches	Conservative	Mismatches	Indels	Gaps
1	281	GNMSSSLF 289	407 AA, 43786 MW, 651002040f68B3D CR664;	71.8%; 55.6%;	28; 58;	5;	2;	2;	0;	0;
<p>RESULT 13</p> <p>MENE_HA3IN</p> <p>ID MENE_HA3IN STANDARD; PRT; 452 AA.</p> <p>AC P44565;</p> <p>DT 01-NOV-1995 (Rel. 32, Created)</p> <p>DT 01-NOV-1995 (Rel. 32, Last sequence update)</p> <p>DT 28-FEB-2003 (Rel. 41, Last annotation update)</p> <p>DE O-succinylbenzoic acid-CoA ligase (EC 6.2.1.26) (OSB-CoA synthetase)</p> <p>DE (O-succinylbenzoate-CoA synthase).</p> <p>GN MENE OR H10194.</p> <p>OS Haemophilus influenzae.</p> <p>OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;</p> <p>OC Pasteurellaceae; Haemophilus.</p> <p>OX NCBI_TaxID=727;</p> <p>RN 1]</p> <p>RP SEQUENCE FROM N.A.</p> <p>RC STRAIN=Rd / KM20 / ATCC 51907;</p> <p>RX MBDLID=95350630; PubMed=7542800;</p> <p>RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,</p> <p>RA Koriavase A.R., Sutton G., Tomb J.-F., Dougherty B.A., Merrick J.M.,</p> <p>RA McKenney K., Sutcliffe G., Fitzhugh W., Fields C.A., Gocayne J.D.,</p> <p>RA Scott J.D., Shrivley R., Liu L.-T., Glodek A., Kelley J.M.,</p> <p>RA Weidman J.F., Phillips C.A., Spriggs T., Hedlorn E., Cotton M.D.,</p> <p>RA Uetendack T.R., Hanna W.C., Nguyen D.T., Saudek D.M., Brandon R.C.,</p> <p>RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,</p> <p>RA Grehn C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,</p>										

RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512 (1995).
 CC -1- FUNCTION: O-SUCCINYLBENZOIC ACID (OSB) TO O-SUCCINYLBENZOYL-COA
 CC (OSB-COA).
 CC -1- CATALYTIC ACTIVITY: ATP + O-succinylbenzoate + CoA = AMP +
 CC diphosphate + O-succinylbenzoyl-CoA.
 CC -1- PATHWAY: Menagunone biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U32704; AAC21863.1; -.
 CC PIR; F64053; F64053.
 DR TIGR; H10194; -.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00501; AMP-binding; 1.
 DR PROSITE; PS00455; AMP-BINDING; 1.
 KW Menagunone biosynthesis; Ligase; Complete proteome.
 SQ SEQUENCE 452 AA; 51394 MW; 3EFA6456AC6F7CBF CRC64;
 Query Match 71.8%; Score 28; DB 1; Length 452;
 Best Local Similarity 75.0%; Pred. No. 64;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GUNAXSAL 8
 DB 47 GUNASAV 54

RESULT 14
 ID VPAP_HSV11 STANDARD; PRT; 488 AA.
 AC P10226;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE DNA polymerase processivity factor (Polymerase accessory protein)
 DE (PAP) (DNA-binding protein U142).
 GN U142.
 OS Herpes simplex virus (type 1 / strain 17).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 CX NCBI_TaxID=10299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88274327; PubMed=2839594;
 RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
 RA McNab D., Perry L.J., Scott J.E., Taylor P.;
 RT "The complete DNA sequence of the long unique region in the genome of
 RT herpes simplex virus type 1.";
 RL J. Gen. Virol. 69:1531-1574 (1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88091053; PubMed=2826807;
 RA McGeoch D.J., Dalrymple M.A., Dolan A., McNab D., Perry L.J.,
 RA Taylor P., Chailberg M.D.;
 RT "Structures of herpes simplex virus type 1 genes required for
 RT replication of virus DNA.";
 RL J. Virol. 62:444-453 (1988).
 RN [3]
 RP FUNCTION.
 RX MEDLINE=91056567; PubMed=2173776;
 RA Gottlieb J., Marcy A.I., Coen D.M., Chailberg M.D.;
 RT "The herpes simplex virus type 1 UL42 gene product: a subunit of DNA

RT polymerase that functions to increase processivity.";
 RL J. Virol. 64:5976-5987 (1990).
 CC -1- FUNCTION: ACCESSORY SUBUNIT OF THE DNA POLYMERASE THAT ACTS TO
 CC INCREASE THE PROCESSIVITY OF POLYMERIZATION.
 CC -1- SIMILARITY: TO OTHER HERPESVIRUSES POLYMERASE ACCESSORY PROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X14112; CAA32305.1; -.
 CC PIR; D28890; WMBE42.
 DR PIR; IDML; 15-MAR-00.
 DR InterPro; IPR003202; U142.
 DR Pfam; PF02282; U142; 2.
 KW DNA-binding; DNA replication; 3D-structure.
 SQ SEQUENCE 488 AA; 51159 MW; 432974563DFOA81B CRC64;
 Query Match 71.8%; Score 28; DB 1; Length 488;
 Best Local Similarity 66.7%; Pred. No. 70;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GUNAXSALF 9
 DB 305 GUNASAVF 313

RESULT 15
 ID CHOD_STRSQ STANDARD; PRT; 546 AA.
 AC P12676;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cholesterol oxidase precursor (EC 1.1.3.6) (CHOD).
 GN CHOA.
 OS Streptomyces sp. (strain SA-COO).
 OC Bacteria; Actinobacteria; Actinobacteriales;
 OC Streptomycetaceae; Streptomycetaceae; Streptomyces.
 CX NCBI_TaxID=1931;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6913081; PubMed=2914858;
 RA Ishizaki T., Hirayama N., Shinikawa H., Nimi O., Murooka Y.;
 RA "Nucleotide sequence of the gene for cholesterol oxidase from a
 RT Streptomyces sp.";
 RL J. Bacteriol. 171:596-601 (1989).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS), AND MUTAGENESIS.
 RX MEDLINE=99211873; PubMed=10194345;
 RA Yue Q.K., Kass I.O., Sampson N.S., Vrieland A.;
 RT "Crystal structure determination of cholesterol oxidase from
 RT Streptomyces and structural characterization of key active site
 RT mutants.";
 RL Biochemistry 38:4277-4286 (1999).
 CC -1- CATALYTIC ACTIVITY: Cholesterol + O(2) = cholest-4-en-3-one +
 CC H2O(2).
 CC -1- COFACTOR: FAD
 CC -1- PATHWAY: CHOLESTEROL METABOLISM.
 CC -1- SUBUNIT: Monomer.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE GMC OXIDOREDUCTASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M1939; AAA26719.1; -

DR PIR; A32260; A32260.

DR PDB; 1B4V; 06-JAN-99.

DR PDB; 1B8S; 09-FEB-99.

DR PDB; 1CBO; 10-MAR-99.

DR PDB; 1CC2; 11-MAR-99.

DR PDB; 1LH; 28-DEC-01.

DR PDB; 1WXT; 25-FEB-03.

DR InterPro; IPR00172; GMC_oxred.

DR InterPro; IPR006311; Tat_C; 1.

DR Pfam; PF05199; GMC_oxred_C; 1.

DR TIGRfams; TIGR01409; Tat_signal_seq; 1.

DR PROSITE; PS00623; GMC_OXRED_1; 1.

DR PROSITE; PS00624; GMC_OXRED_2; FALSE NEG.

KW Oxidoreductase; Signal; Flavoprotein; FAD; 3D-structure.

FT SIGNAL 1 42

FT CHAIN 43 546

FT NP_BIND 54 70

FT ACT_SITE 398 398

FT ACT_SITE 484 484

FT MUTAGEN 398 398

FT MUTAGEN 484 484

FT STRAND 47 53

FT HELIX 57 68

FT TURN 69 70

FT STRAND 73 77

FT TURN 87 88

FT TURN 95 96

FT HELIX 100 102

FT STRAND 104 104

FT STRAND 108 108

FT TURN 113 114

FT HELIX 116 119

FT HELIX 120 122

FT STRAND 125 125

FT STRAND 130 130

FT STRAND 133 137

FT STRAND 142 146

FT TURN 149 150

FT HELIX 151 154

FT TURN 155 155

FT STRAND 159 159

FT HELIX 165 171

FT TURN 173 174

FT HELIX 177 182

FT TURN 183 183

FT HELIX 184 192

FT TURN 193 193

FT STRAND 195 195

FT HELIX 199 204

FT HELIX 206 208

FT TURN 209 220

FT TURN 221 222

FT STRAND 225 227

FT STRAND 230 230

FT STRAND 232 232

FT HELIX 234 241

FT TURN 242 243

FT HELIX 249 251

FT TURN 252 252

FT TURN 255 256

FT STRAND 262 262

FT TURN 265 268

FT HELIX 269 275

FT TURN 276 277

FT STRAND 279 283

FT STRAND 285 292

FT TURN 294 295

FT STRAND 298 305

FT TURN 307 308

FT STRAND 311 324

FT HELIX 327 340

FT TURN 341 342

FT TURN 345 346

FT TURN 349 352

FT STRAND 355 356

FT TURN 358 359

FT STRAND 360 366

FT TURN 369 370

FT STRAND 383 387

FT TURN 392 393

FT STRAND 395 400

FT STRAND 410 417

FT STRAND 423 428

FT TURN 429 432

FT STRAND 433 437

FT HELIX 440 443

FT HELIX 444 461

FT TURN 462 462

FT STRAND 464 465

FT STRAND 477 478

FT STRAND 481 482

FT STRAND 486 486

FT TURN 491 494

FT STRAND 495 495

FT TURN 497 498

FT STRAND 500 501

FT TURN 502 503

FT STRAND 507 509

FT HELIX 512 514

FT TURN 523 523

FT HELIX 524 541

FT TURN 542 542

SQ SEQUENCE 546 AA; 56993 MW; EF22A1FE58A6BD21 CRC64;

Query Match 71.8%; Score 28; DB 1; Length 546;
 Best Local Similarity 75.0%; Pred. No. 79;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 VNAAKALF 9
 DB 447 VNAAKALF 454

Search completed: January 29, 2004, 14:53:11
 Job time : 6.82353 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:46:19 ; Search time 26.6676 Seconds

(without alignments)
86,441 Million cell updates/sec

Title: US-10-032-950-6

Perfect score: 39

Sequence: 1 GYNAXSALP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeophages:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	84.6	47	16	O33586
2	32	82.1	46	2	O33586 staphylococ
3	32	82.1	189	11	O92279
4	32	82.1	406	11	O8VH34 mus musculu
5	32	82.1	407	11	O9DC13 mus musculu
6	32	82.1	452	16	O8RGJ3 fusobacteri
7	31	79.5	308	16	O8DH49 synchococc
8	31	79.5	333	16	O34933 bacillus su
9	31	79.5	478	16	O8F8N1 leptospira
10	31	79.5	593	16	O9PKX0 chlamydia m
11	31	79.5	810	11	O88930 rattus norv
12	31	79.5	1049	3	O13344 emericella
13	30	76.9	234	2	O50375 mycobacteri
14	30	76.9	298	5	O93727 caenorhabdi
15	30	76.9	337	16	O8XW78 clostridium
16	30	76.9	344	16	O8ETK0 oceanobacil

17	30	76.9	361	16	O8UK10	O8UK10 agrobacteri
18	30	76.9	388	2	O9RG47	O9RG47 streptococ
19	30	76.9	549	16	O25619	O25619 helicobacte
20	30	76.9	1795	2	O93UY8	O93UY8 agrobacteri
21	29	74.4	188	16	O9K871	O9K871 bacillus ha
22	29	74.4	196	16	O9CXC1	O9CXC1 pasteurella
23	29	74.4	250	16	O8ZBFI	O8ZBFI yersinia pe
24	29	74.4	255	16	O92DX7	O92DX7 listeria in
25	29	74.4	282	16	O8BSM4	O8BSM4 streptococ
26	29	74.4	282	16	O8DX3	O8DX3 streptococ
27	29	74.4	300	16	O8PE17	O8PE17 xanthomonas
28	29	74.4	305	16	O8PHE1	O8PHE1 xanthomonas
29	29	74.4	327	5	O9BS58	O9BS58 caenorhabdi
30	29	74.4	327	5	O9BS58	O9BS58 caenorhabdi
31	29	74.4	334	9	O9ZKX2	O9ZKX2 mycobacteri
32	29	74.4	336	10	O944F5	O944F5 oryza sativ
33	29	74.4	437	2	O8KX72	O8KX72 aeromonas h
34	29	74.4	445	16	O8E8Y0	O8E8Y0 shewanella
35	29	74.4	461	16	O92R90	O92R90 rhizobium m
36	29	74.4	462	17	O8TPE0	O8TPE0 methanosarc
37	29	74.4	468	2	O8KU05	O8KU05 rhizobium l
38	29	74.4	468	16	O98A92	O98A92 rhizobium l
39	29	74.4	469	17	O8PRB8	O8PRB8 methanosarc
40	29	74.4	480	16	O98J92	O98J92 rhizobium l
41	29	74.4	605	16	O84063	O84063 chlamydia t
42	29	74.4	895	10	O9SC24	O9SC24 arabidopsis
43	29	74.4	895	10	O94C93	O94C93 arabidopsis
44	29	74.4	1016	5	O17465	O17465 plodia inte
45	29	74.4	1016	5	O17484	O17484 plodia inte

ALIGNMENTS

RESULT 1

ID O33586 PRELIMINARY; PRT; 47 AA.

AC O33586;

DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE AGRD (AGRD protein)

GN AGRD OR SAV2037 OR SA1942.1 OR SAS066.

OS Staphylococcus aureus (strain Mu50 / ATCC 700699),

OS Staphylococcus aureus (strain N315), and

OS Staphylococcus aureus.

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI_TaxId=158878, 158879, 1280;

RN (1)

RP SEQUENCE FROM N.A.

RC SPECIES=S. aureus; STRAIN=SA502A;

RX MEDLINE=97342847; PubMed=197262;

RA Ji G., Beavis R., Novick R.P.;

RT "Bacterial interference caused by autoinducing peptide variants.";

RL Science 276:2027-2030 (1997).

RN (2)

RP SEQUENCE FROM N.A.

RC SPECIES=S. aureus (strain Mu50), and S. aureus (strain N315);

RX MEDLINE=21311952; PubMed=11418146;

RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,

RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Ikeno T., Ito T.,

RA Kanomori M., Matsumaru H., Katayama A., Murakami H., Hoshiyama A.,

RA Watanabe-U. Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,

RA Sekimizu K., Hirakawa H., Kunita S., Goto S., Yabuzaki J.,

RA Kanehisa M., Yamashita A., Osima K., Furuya K., Yoshino C., Shiba T.,

RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;

RT "Whole genome sequencing of methicillin-resistant Staphylococcus

RT aureus.";

RL Lancet 357:1225-1240 (2001).

DR EMBL; AF001782; AB63265.1; -

DR EMBL; AP003364; BAB58199.1; -

DR EMBL; AP003135; BAB3124.1; -

Complete proteome.

SQ SEQUENCE 47 AA; 5149 MM; 22893764DF54BFA7 CRC64;

Query Match
Best Local Similarity 84.6%; Score 33; DB 16; Length 47;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
DB 24 GVNACSSLF 32

RESULT 2

Q9FI05 PRELIMINARY; PRT; 46 AA.

AC Q9FI05; 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
GN AgRD.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_Taxid=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN#125;
RX MEDLINE=21137858; PubMed=11240104;
RA Takeuchi S., Maeda T., Hashimoto N., Imaiuni K., Kaidoh T., Hayakawa Y.;
RT "Variation of the agr locus in Staphylococcus aureus isolates from cows with mastitis";
RL Vet. Microbiol. 79:267-274 (2001).
DR EMBL; AB043554; BAB18547.1; 404714CDC4BFA77E CRC64;
SQ SEQUENCE 46 AA; 5050 MM; 404714CDC4BFA77E CRC64;

Query Match
Best Local Similarity 82.1%; Score 32; DB 2; Length 46;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
DB 24 GVNACSSLF 32

RESULT 3

Q922T9 PRELIMINARY; PRT; 189 AA.

AC Q922T9; 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DE 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Similar to lysosomal membrane glycoprotein 1 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RU Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006785; AAH06785.1; -.
DR InterPro; IPR002000; Lamp.
DR Pfam; PF01299; Lamp; 1.
DR PRINTS; PS00336; LYSASOCTDMP.
DR PROSITE; PS00310; LAMP_1; 1.
DR PROSITE; PS00311; LAMP_2; 1.
FT NON TER 1
SQ SEQUENCE 189 AA; 20456 MM; 71F16D5BBA4066FA CRC64;

Query Match
Best Local Similarity 82.1%; Score 32; DB 11; Length 189;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
DB 63 GVNACSSLF 71

RESULT 4

Q9VH34 PRELIMINARY; PRT; 406 AA.

AC Q9VH34; 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DE 01-OCT-2002 (TREMblrel. 22, Last annotation update)
GN LAMP-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Ono K., Han J.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY069968; AAL58070.1; -.
DR InterPro; IPR002000; Lamp.
DR Pfam; PF01299; Lamp; 1.
DR PRINTS; PS00336; LYSASOCTDMP.
DR PROSITE; PS00310; LAMP_1; 2.
DR PROSITE; PS00311; LAMP_2; 1.
DR PROSITE; PS00311; LAMP_2; 1.
SQ SEQUENCE 406 AA; 43879 MM; C1BD373548ADFA85 CRC64;

Query Match
Best Local Similarity 82.1%; Score 32; DB 11; Length 406;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
DB 280 GVNACSSLF 288

RESULT 5

Q9DC13 PRELIMINARY; PRT; 407 AA.

AC Q9DC13; 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DE 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Lysosomal membrane glycoprotein 1.
GN LAMP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085650; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Toh M., Ishii Y., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamazaki I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Glass C., King B., Kochiwa H., Kuhl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schmitt L.W., Staudili F., Suzuki K., Tomita M., Wagner L., Watanabe T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boilelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Shoshenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Whiting L., Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S., Hayashizaki Y.;


```

RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL AK004637; BAB23428.1; -.
DR MGI MG1:96745; Lamp1.
DR InterPro: IPR002000; Lamp.
DR Pfam: PF01299; Lamp; 1.
DR PRINTS: PRO0336; LYSAASOCTMP.
DR PROSITE: PS00310; Lamp_1; 2.
DR PROSITE: PS00311; Lamp_2; 1.
SQ SEQUENCE 407 AA; 43336 MW; 696D0C79F627DA84 CRC64;

Query Match
Best Local Similarity 82.1%; Score 32; DB 11; Length 407;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYMAXSALF 9
DB 281 GYMAXSALF 289

RESULT 6
Q8RGL3 PRELIMINARY; PRT; 452 AA.
AC Q8RGL3;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Xaa-His dipeptidase (EC 3.4.13.3).
GN FN0278.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteriia; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=2186394; PubMed=11889109;
RA Kaparatia V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grecklin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,
RA Larsen N., D'Souza M., Malinas T., Pusch G., Haselkorn R.,
RA Fomstein M., Kyriades N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2016(2002).
DR EMBL AE010540; AAL94484.1; -.
DR InterPro: IPR001261; ARGE_DAPE_CPG2.
DR InterPro: IPR002933; Peptidase_M20.
DR Pfam: PF01546; Peptidase_M20; 1.
DR PROSITE: PS00758; ARGE_DAPE_CPG2_1; 1.
KW Hydroxylase; Dipeptidase; Complete proteome.
SQ SEQUENCE 452 AA; 50314 MW; ED12B150B8FA51E CRC64;

Query Match
Best Local Similarity 77.8%; Score 33; DB 16; Length 452;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYMAXSALF 9
DB 260 GYMAXSALF 268

RESULT 7
Q8DH49 PRELIMINARY; PRT; 308 AA.
AC Q8DH49;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE GlyceroL-3-phosphate dehydrogenase.
GN TL2111.
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.

```

```

OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=2222544; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Kato H., Sasamoto S.,
RA Matsumoto A., Iizuchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
Thermosynechococcus elongatus BP-1."
RL DNA Res. 9:123-130(2002).
DR EMBL: AP005376; BAC09663.1; -.
KW Complete proteome.
SQ SEQUENCE 308 AA; 32597 MW; 3B34BD2DDB59CE55 CRC64;

Query Match
Best Local Similarity 79.5%; Score 31; DB 16; Length 308;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYMAXSALF 8
DB 179 GYMAXSALF 186

RESULT 8
Q34933 PRELIMINARY; PRT; 333 AA.
AC Q34933;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE YFMD protein.
GN YFMD.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9394377;
RA Kunitz F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boutsier L., Brans A., Braun M., Brixnell S.C., Bron S.,
RA Brouillet S., Brysch C.V., Caldwell J., Capuano V.J., Carter N.M.,
RA Choi S.K., Codani J.J., Comerion I.F., Cummings N.J., Daniel R.A.,
RA Denicof F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Evington J., Fabre C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goiteau A., Gollightly E.J., Grand G.,
RA Guileppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulle M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kunano M.,
RA Kurita K., Lapidis A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Maeda S., Meisel D., Nakai S., Noback M.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portelle B., Portelle B., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche B., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scifone F.,
RA Sekiguchi J., Sekowaka A., Seror S.V., Seror P., Shin B.S., Solde B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terstra P., Tognoni A.,
RA Toato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambler R., Wandler E., Wandler H., Weitenberger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumschein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the gram-positive bacterium Bacillus
subtilis."
RL Nature 390:249-256(1997).
[2]

```

RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AC327;
 RA MEDLINE=97417488; PubMed=9272861;
 RX Yamamoto H., Uchiyama S., Nigroho F.A., Sekiguchi J.;
 RT Cloning and sequencing of a 35.7 kb in the 70 degree-73 degree region
 of the Bacillus subtilis genome reveal genes for a new two-component
 RT system, three spore germination proteins, an iron uptake system and a
 RT general stress response protein."
 RL Gene 194191-199(1997).
 DR EMBL; Z99106; CAB12580.1; -;
 DR EMBL; D86417; BAA22318.1; -;
 DR InterPro; IPR000522; FecCD.
 DR Pfam; PF01032; FecCD; 1.
 DR Prodom; PD001557; FecCD; 1.
 KW Complete proteome.
 SQ SEQUENCE 333 AA; 35098 MW; 9F4BE91B0BE6761 CRC64;

Query Match 79.5%; Score 31; DB 16; Length 333;
 Best Local Similarity 66.7%; Pred. No. 82;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GVNAXSALF 9
 DB 99 GVNAGASLF 107

RESULT 9
 O8F8N1 PRELIMINARY; PRT; 493 AA.
 AC O8F8N1; -
 DT 01-MAR-2003 (TEMBLrel. 23, Created)
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
 DE Amino acid transporter.
 GN LATI OR LA0524.
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
 RA Ren S.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB011238; AA047723.1; -;
 KW Complete proteome.
 SQ SEQUENCE 493 AA; 54728 MW; 3AA3859D1FE3C24E CRC64;

Query Match 79.5%; Score 31; DB 16; Length 493;
 Best Local Similarity 66.7%; Pred. No. 12+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVNAXSALF 9
 DB 292 GITASALF 300

RESULT 10
 O9PKY0 PRELIMINARY; PRT; 578 AA.
 AC O9PKY0;
 DT 01-OCT-2000 (TEMBLrel. 15, Created)
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE Protein export protein, FHIPF family.
 GN TC0330.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83560;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MOPn / N1G9;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Ulfberg T., Berry K., Bass S.,
 RA Liner K., Weidman J., Knorr H., Craven B., Bowman C., Dooson R.,
 RA Eisen J., Fraser C.M.,
 RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AR39."
 RL Nucleic Acids Res. 28:1397-1406(2000).
 DR EMBL; AF002301; AA73547.1; -;
 DR TIGR; TC0330; -;
 DR InterPro; IPR001712; Bact_FHIPF.
 DR Pfam; PF00771; FHIPF; 1.
 DR PRINTS; PR00949; TYPE3IMAPROT.
 KW Complete proteome.
 SQ SEQUENCE 578 AA; 64153 MW; BAC9C4852DBE7F5C CRC64;

Query Match 79.5%; Score 31; DB 16; Length 578;
 Best Local Similarity 66.7%; Pred. No. 14+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GVNAXSALF 9
 DB 199 GVNLSALF 207

RESULT 11
 O88930 PRELIMINARY; PRT; 810 AA.
 AC O88930;
 DT 01-NOV-1998 (TEMBLrel. 08, Created)
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
 DE Aryl hydrocarbon receptor.
 GN AHR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Han/Wistar;
 RX MEDLINE=98325247; PubMed=9658193;
 RA Pohjanvirta R., Wong J.M.Y., Li W., Harper P.A., Tuomisto J.,
 RA Okey A.B.;
 RT "Point mutation in intron sequence causes altered carboxyl-terminal
 RT structure in the aryl hydrocarbon receptor of the most 2,3,7,8-
 RT tetrahydrodibenzo-p-dioxin-resistant rat strain."
 RL Mol. Pharmacol. 54:86-93(1998).
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 TRANSCRIPTION FACTORS.
 DR EMBL; AF082124; AAC3168.1; -;
 DR InterPro; IPR01092; HLH_basic.
 DR InterPro; IPR001610; PAC.
 DR InterPro; IPR00014; PAS_domain.
 DR Pfam; PF00010; HLH; 1.
 DR Pfam; PF00785; PAC; 1.
 DR Pfam; PF00989; PAS; 1.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00385; PAC; 1.
 DR SMART; SM00391; PAS; 2.
 DR PROSITE; PS00038; HLH_1; 1.
 DR PROSITE; PS00888; HLH_2; 1.
 DR PROSITE; PS50112; PAS; 1.
 KW Receptor.
 SQ SEQUENCE 810 AA; 91522 MW; DIA99BAA241C8D1 CRC64;

Query Match 79.5%; Score 31; DB 11; Length 810;
 Best Local Similarity 55.6%; Pred. No. 2+02; 1; Indels 0; Gaps 0;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
 DB 758 GINSQSALF 766

RESULT 12

ID 013344 PRELIMINARY; PRT; 1049 AA.
 AC 013344;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JUN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Alpha-mannosidase.
 GN MSD2.
 OS *Emicella nidulans* (*Aspergillus nidulans*).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiaceae; Trichocomaceae; *Emicella*.
 OX NCBI_TaxID=162425;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T222;
 RA Eades C.J., Gilbert A., Goodman C.D., Hintz W.E.;
 RT "Identification and analysis of a class 2 alpha-mannosidase in
 RT *Aspergillus nidulans*.";
 RL Glycobiology 0:0-0(1997).
 DR EMBL; AF016850; AAB70514.1; -;
 DR InterPro; IPR000602; Glyco_hydro_38.
 DR InterPro; IPR006061; SBP_dom1.
 DR Pfam; PF01074; Glyco_hydro_38; 1.
 DR PROSITE; PS01037; SBP_BACTERIAL_1; 1.
 SQ SEQUENCE 1049 AA; 118684 MW; 730627BCA8A999B CRC64;

Query Match 79.5%; Score 31; DB 3; Length 1049;
 Best Local Similarity 87.5%; Pred. No. 2.6e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VNXASALF 9
 DB 979 VNXASALF 986

RESULT 13

ID 050375 PRELIMINARY; PRT; 234 AA.
 AC 050375;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE FXUC (Fragment).
 GN FXUC.
 OS *Mycobacterium smegmatis*.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; *Mycobacterium*.
 OX NCBI_TaxID=1772;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=mc6;
 RX MEDLINE=95191405; PubMed=7885234;
 RA Fies E.H.;
 RT "Identification of genes involved in the sequestration of iron in
 RT *Mycobacterium*: The ferric exochelatin biosynthetic and uptake pathways.";
 RL Mcl. Microbiol. 14:557-569(1994).
 DR EMBL; U10425; AAC43258.1; -;
 DR InterPro; IPR000522; FeCCD.
 DR Pfam; PF01032; FeCCD; 1.
 RT NON TER 1
 SQ SEQUENCE 234 AA; 24159 MW; D4A0F2E75486F861 CRC64;

Query Match 76.3%; Score 30; DB 2; Length 234;
 Best Local Similarity 55.6%; Pred. No. 97;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
 DB 37 GINSQSALF 45

RESULT 14

ID 093727 PRELIMINARY; PRT; 298 AA.
 AC 093727;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE F45H11.1 protein.
 GN F45H11.1.
 OS *Caenorhabditis elegans*.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; *Caenorhabditis*.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kelly F.F.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RP MEDLINE=99069613; PubMed=8851916;
 RA none;
 RT "Genome sequence of the nematode *C. elegans*: A platform for
 RT investigating biology.";
 RT Science 282:2012-2018(1998).
 DR EMBL; Z78420; CAB01709.1; -;
 DR WormPeP; F45H11.1; CE16055.
 DR InterPro; IPR007087; ZnF_C2H2.
 DR Pfam; PF00096; zf-C2H2; 1.
 DR SMART; SM00355; ZNF_C2H2; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
 DR Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 298 AA; 31721 MW; F83B5FE50757527 CRC64;

Query Match 76.9%; Score 30; DB 5; Length 298;
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSALF 8
 DB 159 GINAQSALF 166

RESULT 15

ID 08XM78 PRELIMINARY; PRT; 337 AA.
 AC 08XM78;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Probable ferrichrome ABC transporter.
 GN FHUG OR CPE0812.
 OS *Clostridium perfringens*.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=13 / Type A;
 RX PubMed=11792842;
 RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
 RT "Complete genome sequence of *Clostridium perfringens*, an anaerobic
 RT flesh-eater.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 RL EMBL; AP003188; BAB80518.1; -;
 DR InterPro; IPR000522; FeCCD.
 DR InterPro; IPR001993; Mitoch_carrier.

DR Pfam; PF01032; RecCD; 1.
 DR ProDom; PD001557; RecCD; 1.
 DR PROSITE; PS00215; MITOCH_CARRIER; 1.
 KM Complete proteome.
 SQ SEQUENCE 337 AA; 36437 MW; 565F813B7BCB913 CRC64;

Query Match 76.9%; Score 30; DB 16; Length 337;
 Best Local Similarity 55.6%; Pred. No. 1.4e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYNAXSALF 9
 | : : : |
 Db 96 GINSGALF 104

Search completed: January 29, 2004, 14:56:50
 Job time : 27.8676 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:52:19 ; Search time 11.1176 Seconds
(without alignments)
34.252 Million cell updates/sec

Title: US-10-032-950-6

Perfect score: 39

Sequence: 1 GVNAXSALF 9

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/ptodata/1/aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/aa/5A.COMB.pep:*
4: /cgn2_6/ptodata/1/aa/5B.COMB.pep:*
5: /cgn2_6/ptodata/1/aa/PTCUS.COMB.pep:*
6: /cgn2_6/ptodata/1/aa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	94.9	9	US-09-339-511-6	Sequence 6, Appl
2	34	87.2	9	US-09-339-511-1	Sequence 1, Appl
3	33	84.6	9	US-08-861-476C-3	Sequence 3, Appl
4	33	84.6	4	US-08-861-476C-6	Sequence 6, Appl
5	31	79.5	9	US-09-339-511-5	Sequence 5, Appl
6	30	76.9	9	US-09-339-511-2	Sequence 2, Appl
7	29	74.4	9	US-09-339-511-7	Sequence 7, Appl
8	28	71.8	9	US-09-339-511-4	Sequence 4, Appl
9	28	71.8	322	US-09-312-283C-141	Sequence 141, App
10	28	71.8	322	US-09-312-283C-141	Sequence 141, App
11	28	71.8	341	US-09-145-624-2	Sequence 2, Appl
12	28	71.8	431	US-09-188-930-341	Sequence 341, App
13	28	71.8	431	US-09-312-283C-341	Sequence 341, App
14	28	71.8	488	5223391-9	Patent No. 5223391
15	28	71.8	1380	US-09-252-991A-25772	Sequence 2572, A
16	27	69.2	69	US-09-134-001C-4106	Sequence 4106, Ap
17	27	69.2	217	US-09-252-991A-29222	Sequence 29222, A
18	27	69.2	221	US-08-393-058-1	Sequence 1, Appl
19	27	69.2	350	US-09-252-991A-32358	Sequence 3358, A
20	27	69.2	368	US-09-134-001C-3909	Sequence 3909, Ap
21	27	69.2	407	US-08-776-585-3	Sequence 3, Appl
22	27	69.2	407	US-08-986-659B-10	Sequence 10, Appl
23	27	69.2	410	US-09-594-506-34	Sequence 34, Appl
24	27	69.2	411	US-09-107-532A-7165	Sequence 7165, Ap
25	27	69.2	447	US-09-252-991A-20563	Sequence 20563, A
26	27	69.2	511	US-09-002-285-88	Sequence 88, Appl
27	27	69.2	511	US-09-589-477-88	Sequence 88, Appl

28	27	69.2	536	4	US-08-594-506-32	Sequence 32, Appl
29	27	69.2	746	2	US-08-838-219B-6	Sequence 6, Appl
30	27	69.2	746	3	US-09-233-336A-6	Sequence 6, Appl
31	27	69.2	746	3	US-09-233-752A-6	Sequence 6, Appl
32	27	69.2	746	3	US-09-402-036-6	Sequence 6, Appl
33	27	69.2	746	4	US-09-904-326-6	Sequence 6, Appl
34	27	69.2	789	1	US-08-471-033-29	Sequence 29, Appl
35	27	69.2	789	1	US-08-471-033-32	Sequence 32, Appl
36	27	69.2	789	2	US-08-471-044-29	Sequence 29, Appl
37	27	69.2	789	2	US-08-471-044-32	Sequence 32, Appl
38	27	69.2	789	2	US-08-463-483A-29	Sequence 29, Appl
39	27	69.2	789	2	US-08-463-483A-32	Sequence 32, Appl
40	27	69.2	789	2	US-08-471-046A-29	Sequence 29, Appl
41	27	69.2	789	2	US-08-471-046A-32	Sequence 32, Appl
42	27	69.2	789	2	US-08-470-566B-29	Sequence 29, Appl
43	27	69.2	789	2	US-08-470-566B-32	Sequence 32, Appl
44	27	69.2	789	2	US-08-838-219B-2	Sequence 2, Appl
45	27	69.2	789	2	US-08-838-219B-4	Sequence 4, Appl

ALIGNMENTS

```

RESULT 1
US-09-339-511-6
; Sequence 6, Application US/09339511
; Patent No. 6337385
; GENERAL INFORMATION:
; APPLICANT: Mult, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. 6337385ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 6
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: PEPTIDE
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-6

Query Match          94.9%: Score 37; DB 4; Length 9;
Best Local Similarity 100.0%: Pred. No. 2.5e+05;
Matches          9; Conservative          0; Mismatches          0; Gaps          0;
Db          1 GVNAXSALF 9
          1 GVNAXSALF 9

RESULT 2
US-09-339-511-1
; Sequence 1, Application US/09339511
; Patent No. 6337385
; GENERAL INFORMATION:
; APPLICANT: Mult, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. 6337385ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong

```

TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/09/339,511
CURRENT FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ. ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 1
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-1

Query Match 87.2%; Score 34; DB 4; Length 9;
Best Local Similarity 86.9%; Pred. No. 2.5e+05;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVNAXSALF 9
Db 1 GVNAXSSLF 9

RESULT 3
US-08-861-476C-3
Sequence 3, Application US/08861476C
Patent No. 6447786
GENERAL INFORMATION:
APPLICANT: New York University Medical Center
TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
FILE REFERENCE: 63753/7
CURRENT APPLICATION NUMBER: US/08/861,476C
CURRENT FILING DATE: 1997-05-22
NUMBER OF SEQ. ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 3
LENGTH: 9
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-08-861-476C-3

Query Match 84.6%; Score 33; DB 4; Length 9;
Best Local Similarity 77.8%; Pred. No. 2.5e+05;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GVNAXSALF 9
Db 1 GVNAXSSLF 9

RESULT 4
US-08-861-476C-6
Sequence 6, Application US/08861476C
Patent No. 6447786
GENERAL INFORMATION:
APPLICANT: New York University Medical Center
TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
FILE REFERENCE: 63753/7
CURRENT APPLICATION NUMBER: US/08/861,476C
CURRENT FILING DATE: 1997-05-22
NUMBER OF SEQ. ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 6
LENGTH: 47
TYPE: PRT
ORGANISM: Staphylococcus aureus

US-08-861-476C-6

Query Match 84.6%; Score 33; DB 4; Length 47;
Best Local Similarity 77.8%; Pred. No. 1.1;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GVNAXSALF 9
Db 24 GVNAXSSLF 32

RESULT 5
US-09-339-511-5
Sequence 5, Application US/09339511
Patent No. 6337385
GENERAL INFORMATION:
APPLICANT: Mult. Tom
APPLICANT: Mayville, Patricia
APPLICANT: No. 6337385ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: Ji, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/09/339,511
CURRENT FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ. ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 5
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: peptide
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-5

Query Match 79.5%; Score 31; DB 4; Length 9;
Best Local Similarity 77.8%; Pred. No. 2.5e+05;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVNAXSALF 9
Db 1 GVNAXSSLF 9

RESULT 6
US-09-339-511-2
Sequence 2, Application US/09339511
Patent No. 6337385
GENERAL INFORMATION:
APPLICANT: Mult. Tom
APPLICANT: Mayville, Patricia
APPLICANT: No. 6337385ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: Ji, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/09/339,511
CURRENT FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ. ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 2
LENGTH: 9
TYPE: PRT

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-2

Query Match
Best Local Similarity 76.9%; Score 30; DB 4; Length 9;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GMAXSALF 9
DB 1 GMAXSSLF 9

RESULT 7
US-09-339-511-7
; Sequence 7, Application US/09339511
; Patent No. 6337385
; GENERAL INFORMATION:
; APPLICANT: Muir, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. 6337385ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/09/339,511
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: 60/090,402
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 7
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-7

Query Match
Best Local Similarity 74.4%; Score 29; DB 4; Length 9;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GMAXSALF 9
DB 1 GMAXSSALF 9

RESULT 8
US-09-339-511-4
; Sequence 4, Application US/09339511
; Patent No. 6337385
; GENERAL INFORMATION:
; APPLICANT: Muir, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. 6337385ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/09/339,511
; CURRENT FILING DATE: 1999-06-24
```

```

; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 4
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-4

Query Match
Best Local Similarity 71.8%; Score 28; DB 4; Length 9;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VMAXSALF 9
DB 2 VMAXSSLF 9

RESULT 9
US-09-188-930-141
; Sequence 141, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq For Windows Version 3.0
; SEQ ID NO 141
; LENGTH: 322
; TYPE: PRT
; ORGANISM: mouse
US-09-188-930-141

Query Match
Best Local Similarity 71.8%; Score 28; DB 3; Length 322;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GMAXSALF 9
DB 305 GVNLFSLF 313

RESULT 10
US-09-312-283C-141
; Sequence 141, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
```

NUMBER OF SEQ ID NOS: 425
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 141
LENGTH: 322
TYPE: PRT
ORGANISM: Mouse
US-09-312-283C-141

Query Match 71.8%; Score 28; DB 4; Length 322;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
DB 305 GVNLFSCLP 313

RESULT 11
US-09-149-624-2
Sequence 2, Application US/09149624
Patent No. 6222026
GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
APPLICANT: Kosmatka, Anna L.
APPLICANT: Traini, Christopher M.
TITLE OF INVENTION: gcp
FILE REFERENCE: GML01066
CURRENT APPLICATION NUMBER: US/09/149,624
CURRENT FILING DATE: 1998-09-08
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 341
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-149-624-2

Query Match 71.8%; Score 28; DB 3; Length 341;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSALF 8
DB 95 GVNAXAL 102

RESULT 12
US-09-188-930-341
Sequence 341, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
TITLE OF INVENTION: and Methods For Their Use
FILE REFERENCE: 11000.1011c2
CURRENT APPLICATION NUMBER: US/09/188,930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 341
LENGTH: 431
TYPE: PRT
ORGANISM: Mouse
US-09-188-930-341

Query Match 71.8%; Score 28; DB 3; Length 431;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
DB 305 GVNLFSCLP 313

RESULT 13
US-09-312-283C-341
Sequence 341, Application US/09312283C
Patent No. 6573095
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James G.
APPLICANT: Kumble, Krishanand D.
TITLE OF INVENTION: Compositions Isolated from Skin Cells
TITLE OF INVENTION: and Methods for Their Use
FILE REFERENCE: 11000.1011c2
CURRENT APPLICATION NUMBER: US/09/312,283C
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 425
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 341
LENGTH: 431
TYPE: PRT
ORGANISM: Mouse
US-09-312-283C-341

Query Match 71.8%; Score 28; DB 4; Length 431;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
DB 305 GVNLFSCLP 313

RESULT 14
5223391-9
Patent No. 5223391
APPLICANT: COEN, DONALD M.; DIGARD, PAUL E.
TITLE OF INVENTION: INHIBITORS OF HERPES SIMPLEX VIRUS
REPLICATION
NUMBER OF SEQUENCES: 9
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/482,634
FILING DATE: 21-FEB-1990
SEQ ID NO: 9
LENGTH: 488
5223391-9

Query Match 71.8%; Score 28; DB 6; Length 488;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSALF 9
DB 305 GVNAXSALF 313

RESULT 15
US-09-252-991A-25722
Sequence 25722, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,728
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 25722
 ; LENGTH: 1380
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-25722

Query Match 71.8%; Score 28; DB 4; Length 1380;
 Best Local Similarity 85.7%; Pred No. 6.7e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 NAXSALF 9
 Db 803 NAXSALF 809

Search completed: January 29, 2004, 15:03:22
 Job time : 11.1176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:47:39 ; Search time 25.5441 Seconds
(Without alignments)
73.223 Million cell updates/sec

Title: US-10-032-950-6

Perfect score: 39

Sequence: 1 GVMXSALF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PTC_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PTCUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	94.9	9	US-10-032-950-6	Sequence 6, Appl1
2	34	87.2	9	US-10-032-950-1	Sequence 1, Appl1
3	33	84.6	9	US-10-201-444-3	Sequence 3, Appl1
4	33	84.6	47	US-10-201-444-6	Sequence 6, Appl1
5	31	79.5	9	US-10-032-950-5	Sequence 5, Appl1
6	30	76.9	9	US-10-032-950-2	Sequence 2, Appl1
7	30	76.9	388	US-09-767-041-33	Sequence 33, Appl1
8	30	76.9	508	US-10-369-493-19084	Sequence 19084, A
9	30	76.9	549	US-09-882-227-184	Sequence 184, App
10	29	74.4	9	US-10-032-950-7	Sequence 7, Appl1
11	29	74.4	337	US-10-004-502-12	Sequence 12, Appl1
12	29	74.4	391	US-10-156-761-9222	Sequence 9222, Ap
13	29	74.4	468	US-10-369-493-12194	Sequence 12194, A
14	29	74.4	605	US-09-841-133-574	Sequence 574, A
15	28	71.8	9	US-10-032-950-4	Sequence 4, Appl1

16	28	71.8	122	12	US-09-864-408A-7538	Sequence 7538, Ap
17	28	71.8	144	10	US-09-894-018-129	Sequence 129, App
18	28	71.8	255	15	US-10-083-357-1280	Sequence 1280, Ap
19	28	71.8	322	11	US-09-866-050A-141	Sequence 141, App
20	28	71.8	332	12	US-10-264-237-2080	Sequence 2080, Ap
21	28	71.8	338	9	US-09-815-242-5727	Sequence 5727, Ap
22	28	71.8	341	9	US-09-820-044-2	Sequence 2, Appl1
23	28	71.8	341	9	US-09-815-242-12158	Sequence 12158, A
24	28	71.8	431	11	US-09-866-050A-341	Sequence 341, App
25	28	71.8	431	12	US-10-004-378A-129	Sequence 129, App
26	28	71.8	452	15	US-10-260-877-96	Sequence 96, Appl1
27	28	71.8	491	12	US-10-094-749-2461	Sequence 2461, Ap
28	28	71.8	515	9	US-09-823-356-3	Sequence 3, Appl1
29	28	71.8	639	12	US-10-260-937-39	Sequence 39, Appl1
30	27	70.5	433	12	US-10-369-493-11945	Sequence 11945, A
31	27	69.2	138	9	US-09-864-761-34801	Sequence 34801, A
32	27	69.2	181	12	US-10-004-378A-190	Sequence 190, App
33	27	69.2	248	12	US-10-369-493-14264	Sequence 14264, A
34	27	69.2	248	12	US-10-369-493-14770	Sequence 14770, A
35	27	69.2	257	12	US-10-369-493-11482	Sequence 11482, A
36	27	69.2	280	12	US-10-369-493-981	Sequence 981, App
37	27	69.2	299	12	US-10-017-161-2132	Sequence 2132, Ap
38	27	69.2	299	12	US-10-292-798-1778	Sequence 1778, App
39	27	69.2	330	12	US-10-369-493-2642	Sequence 2642, App
40	27	69.2	340	9	US-09-815-242-11980	Sequence 11980, A
41	27	69.2	425	12	US-10-369-493-376	Sequence 376, App
42	27	69.2	432	9	US-09-799-777-74	Sequence 74, Appl1
43	27	69.2	432	11	US-09-832-129-41	Sequence 41, Appl1
44	27	69.2	432	11	US-09-832-129-54	Sequence 54, Appl1
45	27	69.2	432	12	US-10-024-298A-85	Sequence 85, Appl1

ALIGNMENTS

RESULT 1
US-10-032-950-6
Sequence 6, Application US/10032950
Publication No. US2002007453A1
GENERAL INFORMATION:
APPLICANT: Mult, Tom
APPLICANT: Mayville, Patricia
APPLICANT: No US2002007453A1ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: Ji, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/10/032,950
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-6

Query Match 94.9% Score 37; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GVMXSALF 9
Db 1 GVMXSALF 9

RESULT 2

US-10-032-950-1
; Sequence 1, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mult, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-1

Query Match 87.2%; Score 34; DB 13; Length 9;
Best Local Similarity 88.9%; Pred. No. 7e+05;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GVNAXSALF 9
Db 1 GVNAXSSLF 9

RESULT 3

US-10-201-444-3
; Sequence 3, Application US/10201444
; Publication No. US20030078378A1
; GENERAL INFORMATION:
; APPLICANT: New York University Medical Center
; TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
; FILE REFERENCE: 63753/7
; CURRENT APPLICATION NUMBER: US/10/201,444
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/08/861,476
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-201-444-3

Query Match 84.6%; Score 33; DB 15; Length 9;
Best Local Similarity 77.8%; Pred. No. 7e+05;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GVNAXSALF 9
Db 1 GVNAXSSLF 9

RESULT 4
US-10-201-444-6

; Sequence 6, Application US/10201444
; Publication No. US20030078378A1
; GENERAL INFORMATION:
; APPLICANT: New York University Medical Center
; TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
; FILE REFERENCE: 63753/7
; CURRENT APPLICATION NUMBER: US/10/201,444
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/08/861,476
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-201-444-6

Query Match 84.6%; Score 33; DB 15; Length 47;
Best Local Similarity 77.8%; Pred. No. 1.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GVNAXSALF 9
Db 24 GVNAXSSLF 32

RESULT 5

US-10-032-950-5
; Sequence 5, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mult, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-5

Query Match 79.5%; Score 31; DB 13; Length 9;
Best Local Similarity 77.8%; Pred. No. 7e+05;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GVNAXSALF 9
Db 1 GVNAXSSLF 9

RESULT 6
US-10-032-950-2

; Sequence 2, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mult, Tom

```

; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-2

```

```

Query Match      76.9%; Score 30; DB 13; Length 9;
Best Local Similarity 77.8%; Pred. No. 7e+05;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 GVNAXSALF 9
Db      1 GVNAXSSLF 9

```

```

RESULT 7
US-09-767-041-33
; Sequence 33, Application US/09767041
; Patent No. US20020055168A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Hilda
; TITLE OF INVENTION: STREPTOCOCCUS SUIS VACCINES AND DIAGNOSTIC TESTS
; FILE REFERENCE: 2183-4726
; CURRENT APPLICATION NUMBER: US/09/767,041
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: PCT/NL99/00460
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: EP98202465.5
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: EP98202467.1
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Streptococcus suis
; FEATURES:
; NAME/KEY: misc feature
; OTHER INFORMATION: CPS1H
US-09-767-041-33

```

```

Query Match      76.9%; Score 30; DB 9; Length 388;
Best Local Similarity 66.7%; Pred. No. 95;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 GVNAXSALF 9
Db      335 GVNAXSALF 343

```

```

RESULT 8
US-10-369-493-19084
; Sequence 19084, Application US/10369493

```

```

; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19084
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Anabaena PCC7120
US-10-369-493-19084

```

```

Query Match      76.9%; Score 30; DB 12; Length 508;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 GVNAXSALF 8
Db      156 GVNAXSALF 163

```

```

RESULT 9
US-09-882-227-184
; Sequence 184, Application US/09882227
; Publication No. US20030158396A1
; GENERAL INFORMATION:
; APPLICANT: Kleantsov, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Ooomeen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20030158396A1 Helicobacter Polypeptides in th
; FILE REFERENCE: 06132/047002
; CURRENT APPLICATION NUMBER: US/09/882,227
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 06/902,615
; PRIOR FILING DATE: 1997-07-29
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 184
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-882-227-184

```

```

Query Match      76.9%; Score 30; DB 12; Length 549;
Best Local Similarity 55.6%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 GVNAXSALF 9
Db      406 GVNAXSALF 414

```

```

RESULT 10
US-10-032-950-7
; Sequence 7, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mait, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.

```

```

; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 7
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-7

```

```

Query Match          74.4%; Score 29; DB 13; Length 9;
Best Local Similarity 77.8%; Pred. No. 7e+05; 1; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 1;

```

```

QY 1 GVNAXSALF 9
Db 1 GVNAXSALF 9

```

```

RESULT 11
US-10-004-502-12
; Sequence 12, Application US/10004502
; Publication No. US2003008882A1
; GENERAL INFORMATION:
; APPLICANT: Harwell, Leslie T.
; APPLICANT: Rasghianti, James J
; TITLE OF INVENTION:
; FILE REFERENCE: B1470 US NA
; CURRENT APPLICATION NUMBER: US/10/004,502
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,272
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 12
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Trifolium aestivum
US-10-004-502-12

```

```

Query Match          74.4%; Score 29; DB 15; Length 337;
Best Local Similarity 66.7%; Pred. No. 1.4e+02; 3; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 3;

```

```

QY 1 GVNAXSALF 9
Db 90 GVNAEGCLF 98

```

```

RESULT 12
US-10-156-761-9222
; Sequence 9222, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI

```

```

; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-263
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO: 9222
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9222

```

```

Query Match          74.4%; Score 29; DB 15; Length 391;
Best Local Similarity 77.8%; Pred. No. 1.6e+02; 2; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 2;

```

```

QY 1 GVNAXSALF 9
Db 274 GVAAVSALF 282

```

```

RESULT 13
US-10-369-493-12194
; Sequence 12194, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianheng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO: 12194
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-12194

```

```

Query Match          74.4%; Score 29; DB 12; Length 468;
Best Local Similarity 66.7%; Pred. No. 2e+02; 2; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 2;

```

```

QY 1 GVNAXSALF 9
Db 408 GVRAAALF 416

```

```

RESULT 14
US-09-841-133-574
; Sequence 574, Application US/09841132
; Patent No. US2002061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhalla, Ajay
; APPLICANT: Skeiky, Yaelir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0

```

```

; SEQ ID NO 574
;
; LENGTH: 605
; TYPE: PRT
; ORGANISM: C. Trachomatis D serovar
US-09-841-132-574

```

```

Query Match
Best Local Similarity 74.4%; Score 29; DB 9; Length 605;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 GVNXSALF 9
Db      226 GVNIIALF 234

```

```

RESULT 15
US-10-032-950-4
; Sequence 4, Application US/10032950
; Publication No. US2002007453A1
; GENERAL INFORMATION:
; APPLICANT: Mult. Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US2002007453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-4

```

```

Query Match
Best Local Similarity 71.8%; Score 28; DB 13; Length 9;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 VNAXSALF 9
Db      2 VNAXSSLF 9

```

```

Search completed: January 29, 2004, 15:01:46
Job time : 25.5441 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:44:54 ; Search time 35.0735 Seconds
(without alignments)
40.730 Million cell updates/sec

Title: US-10-032-950-7

Perfect score: 39

Sequence: 1 GVNAXSSAF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	94.9	9	21	AA67857
2	37	94.9	9	23	ABP53546
3	37	94.9	9	23	AAW50905
4	33	84.6	9	21	AA67859
5	33	84.6	9	21	AA67860
6	33	84.6	9	23	AAW50907
7	33	84.6	9	23	AAW51002
8	33	84.6	9	23	AAW51003
9	33	84.6	9	23	AAW51004

10	33	84.6	9	23	ABW07160
11	32	82.1	9	21	AA67851
12	32	82.1	9	21	AA67861
13	32	82.1	9	23	ABP53540
14	32	82.1	9	23	AAW50899
15	32	82.1	9	23	AAW50908
16	32	82.1	9	23	AAW51005
17	32	82.1	9	23	ABW07161
18	31	79.5	9	19	AAW38323
19	31	79.5	9	23	AAW51001
20	31	79.5	9	23	ABW07159
21	31	79.5	9	24	ABW84631
22	31	79.5	47	24	ABW84634
23	31	79.5	924	24	ABW02122
24	30	76.9	163	23	ABW05504
25	30	76.9	167	23	ABW05887
26	30	76.9	167	24	ABP57498
27	30	76.9	416	16	AAW69554
28	30	76.9	416	21	AAW52550
29	30	76.9	417	11	AAW01940
30	30	76.9	614	23	AAW09996
31	30	76.9	614	23	AAW09996
32	30	76.9	945	22	ABW60000
33	29	74.4	9	21	AAW67855
34	29	74.4	9	21	AAW67856
35	29	74.4	9	23	ABP53544
36	29	74.4	9	23	ABP53545
37	29	74.4	9	23	AAW50903
38	29	74.4	9	23	AAW50904
39	29	74.4	49	20	AAW11952
40	29	74.4	273	20	AAW12079
41	29	74.4	273	20	AAW36145
42	29	74.4	273	20	AAW36192
43	29	74.4	273	24	ABW82232
44	29	74.4	273	24	ABW84449
45	29	74.4	273	24	ABW56773

ALIGNMENTS

RESULT 1
AA67857
ID AA67857 standard; peptide; 9 AA.
XX
AC AA67857;
XX
DT 25-APR-2000 (first entry)
XX
DE S. aureus peptide #6 used for bacterial interference.
XX
XX Staphylococcus aureus infection; cyclic peptide; AgrD; agr response;
KW virulence factor; treatment.
XX
XX Staphylococcus aureus.
OS
XX
FH Key
FT Modified-site
FT /note= "N-terminal residue forms bond with C-terminal
FT residue to form a cyclic peptide"
FT
FT Misc-difference 5
FT /label= Unknown
XX
XX
XX WO9967286-A2.
XX
XX PD 29-DEC-1999.
XX
XX PF 24-JUN-1999; 99WO-US14562.
XX
XX PR 24-JUN-1999; 98US-0103438.
XX
XX (UVRQ) UNIV ROCKEFELLER.
XX (UVRV) UNIV NEW YORK STATE.

Peptide-mediated q
S. aureus peptide
Staphylococcus aur
Cyclic peptide SEQ
AgrD-autoinducing
Protected peptide
AgrD2 lactam cycli
Peptide-mediated q
Transcription inh
AgrD2 thiololacton
Peptide-mediated q
S. aureus RN6607 A
S. aureus RN6607 A
S. pneumoniae type
M. tuberculosis an
K. tuberculosis an
Mycobacterium tube
Human lysosomal me
Human lysosomal me
Tumour necrosis fa
Genitourinfectum ha
Haloperoxidase am
Drosophila melanog
S. aureus peptide
S. aureus peptide
Cyclic peptide SEQ
Cyclic peptide SEQ
AgrD-autoinducing
AgrD-autoinducing
Human 5' EST seque
Human gap junction
Human secreted pro
Human bladder can
Lung cancer-associ
Lung cancer-associ

```

XX
PI Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
XX WPI, 2000-147202/13.
DR
XX
XX New cyclic peptides for treating infections with Staphylococcus aureus
PT
XX
XX Claim 9; Page 26; 37pp; English.
XX
XX This sequence represents a cyclic peptide derived from the Staphylococcus
CC aureus AgrD peptide. The invention relates to AgrD derived peptides,
CC a composition containing a peptide and a carrier, and a method for the
CC production of the cyclic peptides. The peptide inhibits the agr response,
CC which is normally associated with the release of virulence factors of
CC Staphylococcus aureus. An AgrD peptide is produced by S. aureus that
CC activates the agr response in strains of a single group, but interferes
CC with this response in strains of different groups. The peptides and
CC composition containing them can be used to treat infections by S. aureus.
XX
SQ Sequence 9 AA;
XX
XX Query Match 94.9%; Score 37; DB 21; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GVNAKSSAF 9
XX |||||
DB 1 GVNAKSSAF 9
XX
XX
XX RESULT 2
XX ABP53546
XX ID ABP53546 standard; peptide; 9 AA.
XX AC
XX ABP53546;
XX
XX 13-DEC-2002 (first entry)
XX DT
XX Cyclic peptide SEQ ID NO:7.
XX DE
XX Cyclic peptide; Staphylococcus aureus; infection; antibacterial;
XX KW agr response inhibitor.
XX
XX Synthetic.
XX OS
XX Key Location/Qualifiers
XX FH Misc-difference 5 /note= "any amino acid"
XX FT
XX US200207453-A1.
XX PN
XX 20-JUN-2002.
XX PD
XX 27-DEC-2001; 2001US-0032950.
XX PF
XX 24-JUN-1998; 98US-090402P.
XX PR 24-JUN-1999; 99US-0339511.
XX
XX (MUIR/) MUIR T W.
XX PA (MAYV/) MAYVILLE P.
XX PA (NOVI/) NOVICK R P.
XX PA (BEAV/) BEAVIS R.
XX PA (JIG/) JI G.
XX
XX Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX WPI, 2002-681366/73.
XX
XX New cyclic peptides, useful for treating Staphylococcus aureus
XX infections
XX
XX Claim 9; Page 10; 18pp; English.

```

```

XX
XX ABP53540 to ABP53547 represent cyclic peptides (I) from the present
CC invention. The present invention also describes a method for treating
CC Staphylococcus aureus infection comprising the administration of a
CC composition comprising (I). (I) has antibacterial activity, and can be
CC used as an agr gene response inhibitor. The peptides are useful for
CC treating S. aureus infections.
XX
XX
XX Sequence 9 AA;
XX
XX Query Match 94.9%; Score 37; DB 23; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GVNAKSSAF 9
XX |||||
DB 1 GVNAKSSAF 9
XX
XX
XX RESULT 3
XX AAM50905
XX ID AAM50905 standard; Peptide; 9 AA.
XX AC
XX AAM50905;
XX
XX 08-MAY-2002 (first entry)
XX DT
XX AgrD-autoinducing cyclic peptide, inhibitor of agr response.
XX DE
XX Staphylococcus aureus; AgrD; agr response; inhibitor; antibiotic;
XX KW antibacterial; infection; therapy; cyclic.
XX
XX Synthetic.
XX OS
XX Key Location/Qualifiers
XX FH Misc-difference 5 /note= "any amino acid"
XX FT
XX Misc-difference 5 /note= "note linked to residue 9 to form cyclic
XX FT peptide"
XX FT Misc-difference 9 /note= "note linked to residue 5 to form cyclic
XX FT peptide"
XX
XX US6337385-B1.
XX PN
XX 08-JAN-2002.
XX PD
XX 24-JUN-1999; 99US-0339511.
XX PF
XX 24-JUN-1998; 98US-090402P.
XX PR
XX (UYRQ ) UNIV ROCKEFELLER.
XX PA (UYNY ) UNIV NEW YORK STATE.
XX
XX Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX WPI, 2002-170774/22.
XX
XX Claim 7; Column 19; 18pp; English.
XX
XX The present sequence is that of a novel synthetic cyclic peptide
CC of the invention that is capable of inhibiting the agr response of
CC Staphylococcus aureus. It is an AgrD-autoinducing peptide, where
CC AgrD is a secreted agr-encoded peptide and where the agr locus
CC controls the synthesis of virulence factor and other extracellular
CC proteins responsible for pathogenicity in S. aureus. Preferred
CC peptides may have the sequence NH2-X(n)-2-X(y)-COOH, with a
XX cyclic bond between the Z residue and COOH other than a thioester

```



```
CC bond, where X is an amino acid, an amino acid analogue, a
CC peptidomimetic or non-amide isostere, Z is a synthetic or a
CC biosynthetic amino acid, n is 0-10 and y is 1-10. The cyclic bond
CC is especially a lactam or lactone bond. The thiolactone
CC structure within native Agard peptides is required for activation
CC of the agr response. Elimination of the thiol ester component of
CC the cyclic ring structure can destroy agr response activating
CC activity while preserving and enhancing inhibitory activity. A
CC claimed method of preparing a cyclic peptide involves assembling
CC a linear peptide chain on to a solid phase resin support;
CC deprotecting the resulting protected assembled peptide; treating the
CC deprotected peptide with neutral buffer for a time sufficient to
CC form the cyclic peptide and cleave the peptide from the support;
CC and recovering the cyclic peptide. The peptide is useful for
CC bacterial interference, especially for the treatment of S. aureus
CC infection.
```

```
XX Sequence      9 AA;
```

```
SQ
```

```
Query Match          94.9%; Score 37; DB 23; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0.
```

```
QY      1 GVNAKSSAF 9
        |||||
Db       1 GVNAKSSAF 9
```

```
RESULT 4
AA67859 ID AA67859 standard; peptide; 9 AA.
XX AC AAY67859;
XX DT 25-APR-2000 (first entry)
XX DE Staphylococcus aureus AgardII derived peptide sequence.
DE Staphylococcus aureus infection; treatment; Agard; agr response;
KW Staphylococcus aureus infection; treatment; Agard; agr response;
KM virulence factor.
XX OS Staphylococcus aureus.
XX PN W09967286-A2.
PN PD 29-DEC-1999.
XX PF 24-JUN-1999; 99WO-US14562.
XX PR 24-JUN-1998; 98US-0103438.
XX PA (UVRQ ) UNIV ROCKEFELLER.
XX PA (UYNY ) UNIV NEW YORK STATE.
PI Muir TM, Mayville P, Novick RP, Ji G, Beavis R;
XX WP1; 2000-147202/13.
DR
PT New cyclic peptides for treating infections with Staphylococcus aureus
PT -
XX Examples; Page 22; 37pp; English.
XX This sequence represents the Staphylococcus aureus AgardII derived
XX peptide. The invention relates to Agard derived cyclic peptides, a
XX composition containing a peptide and a carrier, and a method for the
XX production of the cyclic peptides. The peptide inhibits the agr response,
XX which is normally associated with the release of virulence factors of
XX Staphylococcus aureus. An Agard peptide is produced by S. aureus that
XX activates the agr response in strains of a single group, but interferes
XX with this response in strains of different groups. The peptides and
XX composition containing them can be used to treat infections by S. aureus.
```

```

SQ Sequence 9 AA;
Query Match 84.6%; Score 33; DB 21; Length 9;
Best Local Similarity 77.8%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAKSSAF 9
   ||| |||
Db 1 GVNAKSSLP 9

RESULT 5
AAV67860 standard; peptide; 9 AA.
AAV67860:
AC AAV67860:
AT 25-APR-2000 (first entry)
XX
XX Staphylococcus aureus AgardII derived peptide sequence.
DE
XX Staphylococcus aureus infection; Agard; agr response; treatment;
KM virulence factor.
XX
XX Staphylococcus aureus.
OS
XX WO967286-A2.
XX
XX 29-DEC-1999.
PD
XX
XX 24-JUN-1999; 99WO-US14562.
FP
XX
XX 24-JUN-1998; 98US-0103438.
PR
XX
XX (VYRQ ) UNITV ROCKEFELLER.
PA (UNITV ) UNITV NEW YORK STATE.
XX
XX
XX Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
PI
XX WPI; 2000-147202/13.
DR
XX
XX New cyclic peptides for treating infections with Staphylococcus aureus
PT
XX
XX
XX
XX Examples; Page 22; 37pp; English.
XX
XX This sequence represents the Staphylococcus aureus AgardII derived
CC peptide. The invention relates to Agard derived cyclic peptides, a
CC composition containing a peptide and a carrier, and a method for the
CC production of the cyclic peptides. The peptide inhibits the agr response,
CC which is normally associated with the release of virulence factors of
CC Staphylococcus aureus. An Agard peptide is produced by S. aureus that
CC activates the agr response in strains of a single group, but interferes
CC with this response in strains of different groups. The peptides and
CC composition containing them can be used to treat infections by S. aureus.
XX
XX
XX Sequence 9 AA:
SQ
Query Match 84.6%; Score 33; DB 21; Length 9;
Best Local Similarity 77.8%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAKSSAF 9
   ||| |||
Db 1 GVNAKSSLP 9

RESULT 6
AAV50907
ID AAV50907 standard; Peptide; 9 AA.
AC
XX AAV50907;
XX

```

Query Match	Score	DB	Length
Best Local Similarity	77.8%	Pred. No. 9.3e+05;	
Sequence	9 AA;		

Matches	;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
QY		1 GYNAXSSAF 9							
Db		1 GYNAXSSSLF 9							
RESULT 7									
AA		AA							
ID		AA							
XX		AA							
AC		AA							
XX		AA							
DT		06-MAY-2002							
XX		(first entry)							
DE		Agd2 linear thioester peptide.							
XX									
KW		Agd2; agr response; inhibitor; antibiotic; antibacterial;							
XX		infection; therapy.							
OS		Staphylococcus aureus.							
OS		Synthetic.							
Key		Location/Qualifiers							
FT		Modified-site 9							
FT		/note= "C-terminal thioester"							
XX									
PN		US6337365-B1.							
XX									
PD		08-JAN-2002.							
XX									
PF		24-JUN-1999; 99US-0339511.							
XX									
PR		24-JUN-1998; 98US-090402P.							
XX									
PA		(UNYR) UNIV ROCKEFELLER.							
PA		(UNYR) UNIV NEW YORK STATE.							
PI		Muir TW, Mayville P, Novick RP, Beavis R, Ji G;							
XX									
DR		WPI; 2002-170774/22.							
XX									
PT		Novel synthetic, cyclic Agd2-anticoagulating peptide for bacterial							
PT		interference and for treating Staphylococcus aureus infection in a							
PT		subject -							
XX									
PS		Example 1; Column 9; 18pp; English.							
XX									
CC		The present sequence is that of a novel synthetic Agd2 linear							
CC		thioester peptide. The peptide is derived from the cyclic Agd2							
CC		peptide of Staphylococcus aureus group II. Agd2 is a secreted							
CC		agr-encoded peptide, where the agr locus controls the synthesis of							
CC		virulence factor and other extracellular proteins responsible for							
CC		pathogenicity in S. aureus. The biological activity of the							
CC		synthetic peptide was assayed using cultured S. aureus strains							
CC		containing a beta-lactamase reporter gene fused to the agrP3							
CC		promoter. This allowed activation or inhibition of the agr							
CC		response to be monitored spectrophotometrically. Unlike an Agd2							
CC		thiolactone cyclic peptide (see AAM51001), the present peptide							
CC		was unable to either activate or inhibit the agr response, even							
CC		when added to cultured cells at 10^6 concentrations. The invention							
CC		provides claimed cyclic peptides (see AAM50899-906 and AAM50999)							
CC		and methods for preparing them. The cyclic peptides are useful for							
CC		bacterial interference, especially for the treatment of S. aureus							
CC		infection.							
XX									
XX									

[illegible]

ABB07160
ID ABB07160 standard; peptide; 9 AA.
XX
AC ABB07160;
XX
DT 13-MAR-2002 (first entry)
XX
DE Peptide-mediated quorum sensing inhibitor peptide cyclo-XII.
XX
XX Autoinducer-2; AI-2; antibiotic; antibacterial; dermatological;
KM vulnery; pheromone; agr system; accessory gene regulator; cyclic.
XX
OS Synthetic.
XX
FN WO200185664-A2.
XX
PD 15-NOV-2001.
XX
PF 10-MAY-2001; 2001WO-US15221.
XX
PR 10-MAY-2000; 2000US-203000P.
PR 07-DEC-2000; 2000US-254398P.
XX
PA (UYPR-) UNIV PRINCETON.
PA (QUOR-) QUOREX PHARM INC.
PA (UYTE-) UNIV TECHNOLOGIES INT INC.
XX
PI Basler BL, Dammel CS, Schauder S, Shokat K, Stein J, Surette MG;
XX
XX WPI; 2002-075235/10.
XX
XX Use of autoinducer-2 agonists or antagonists for regulating activity of
PT autoinducer-2 receptor, regulating bacterial growth and pathogenesis,
PT also antibiotic compositions
XX
XX Disclosure; Page 33; 134pp; English.
XX
PS The invention relates to the use of autoinducer-2 (AI-2) agonists or
CC antagonists for regulating activity of autoinducer-2 receptor, regulating
CC bacterial growth and pathogenesis. Synergistic antibiotic compositions
CC comprising inhibitors of the quorum-sensing pathway of a microorganism
CC are also provided. Methods using such AI-2 analogues are useful for
CC treating pathogen-associated disease states. The compounds and antibiotic
CC compositions can be used to inhibit bacterial cell growth and/or biofilm
CC formation on a medical device, particularly for promoting growth of skin
CC graft replacements used in the treatment of burns and ulcers. They may
CC also be used to aid wound repair, and to inhibit bacterial cell growth
CC and biofilm formation in or on products or devices used for personal
CC hygiene. The present sequence represents a inhibitor of peptide-mediated
CC quorum sensing.
CC
XX
SQ Sequence 9 AA;
XX
XX Query Match 84.6%; Score 33; DB 23; Length 9;
Best Local Similarity 77.8%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 11
ID AAY67851
XX AAY67851 standard; peptide; 9 AA.
XX
AC AAY67851;
XX
DT 25-APR-2000 (first entry)
XX
DE S. aureus peptide #1 used for bacterial interference.
XX
KM Staphylococcus aureus infection; cyclic peptide; AgrD; agr response;

KM virulence factor; treatment.
XX
XX Staphylococcus aureus.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 5 /label= Unknown
FT
XX
XX WO967286-A2.
PN
XX
PD 29-DEC-1999.
XX
PF 24-JUN-1999; 99WO-US14562.
XX
PR 24-JUN-1998; 98US-0103438.
XX
XX (UYRQ) UNIV ROCKEFELLER.
PA (UYNY) UNIV NEW YORK STATE.
XX
PI Muller TW, Mayville P, Novick RP, Ji G, Beavis R;
XX
XX WPI; 2000-147202/13.
XX
XX New cyclic peptides for treating infections with Staphylococcus aureus
PT
PT Claim 9; Page 26; 37pp; English.
XX
XX This sequence represents a cyclic peptide derived from the Staphylococcus
CC aureus AgrD peptide. The invention relates to AgrD derived peptides,
CC a composition containing a peptide and a carrier, and a method for the
CC production of the cyclic peptides. The peptide inhibits the agr response,
CC which is normally associated with the release of virulence factors of
CC Staphylococcus aureus. An AgrD peptide is produced by S. aureus that
CC activates the agr response in strains of a single group, but interferes
CC with this response in strains of different groups. The peptides and
CC composition containing them can be used to treat infections by S. aureus.
CC
XX
SQ Sequence 9 AA;
XX
XX Query Match 82.1%; Score 32; DB 21; Length 9;
Best Local Similarity 88.9%; Pred. No. 9.3e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 12
ID AAY67861
XX AAY67861 standard; peptide; 9 AA.
XX
AC AAY67861;
XX
DT 25-APR-2000 (first entry)
XX
DE Staphylococcus aureus AgrDII derived peptide sequence.
XX
KM Staphylococcus aureus infection; AgrD; agr response; treatment;
KM virulence factor.
XX
XX Staphylococcus aureus.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 5 /label= Unknown
FT
XX
XX WO967286-A2.
PN
XX
PD 29-DEC-1999.
XX
PF 24-JUN-1999; 99WO-US14562.

```

XX 24-JUN-1998; 98US-0103438.
XX (UYRQ ) UNIV ROCKEFELLER.
XX PA (UYNY ) UNIV NEW YORK STATE.
XX
XX Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
XX WPI, 2000-147202/13.
XX
XX New cyclic peptides for treating infections with Staphylococcus aureus
PT _
PT Examples; Page 22; 37pp; English.
XX
XX This sequence represents the Staphylococcus aureus AgdII derived
CC peptide. The invention relates to Agd derived cyclic peptides, a
CC composition containing a peptide and a carrier, and a method for the
CC production of the cyclic peptide. The peptide inhibits the agr response,
CC which is normally associated with the release of virulence factors of
CC Staphylococcus aureus. An Agd peptide is produced by S. aureus that
CC activates the agr response in strains of a single group, but interferes
CC with this response in strains of different groups. The peptides and
CC composition containing them can be used to treat infections by S. aureus.
XX
XX Sequence 9 AA;
SQ
XX
XX Query Match 82.1%; Score 32; DB 21; Length 9;
XX Best Local Similarity 88.9%; Pred. No. 9.3e+05;
XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0.
XX
XX 1 GVNAKSSAF 9
XX | | | | | |
XX 1 GVNAKSLTF 9
XX
XX RESULT 13
XX ID ABP53540
XX AC ABP53540 standard; peptide; 9 AA.
XX
XX ABP53540;
XX
XX 13-DEC-2002 (first entry)
XX
XX Cyclic peptide SEQ ID NO:1.
XX
XX Cyclic peptide; Staphylococcus aureus; infection; antibacterial;
XX agr response inhibitor.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Key Misc-difference 5 /note= "any amino acid"
XX
XX US2002077453-A1.
XX
XX 20-JUN-2002.
XX
XX 27-DEC-2001; 2001US-0032850.
XX
XX 24-JUN-1998; 98US-090402P.
XX PR 24-JUN-1999; 99US-0339511.
XX
XX (MUIR/) MUIR T W.
XX (MAYV/) MAYVILLE P.
XX (NOVI/) NOVICK R P.
XX (BEAV/) BEAVIS R.
XX (JTG/) JI G.
XX
XX Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
XX WPI: 2002-681366/73.

```

XX	New cyclic peptides, useful for treating <i>Staphylococcus aureus</i>
FT	infections -
XX	
PS	Claim 9; Page 10; 18pp; English.
XX	
CC	ABP53540 to ABP53547 represent cyclic peptides (I) from the present
CC	invention. The present invention also describes a method for treating
CC	<i>Staphylococcus aureus</i> infection comprising the administration of a
CC	composition comprising (I). (I) has antibacterial activity, and can be
CC	used as an agr gene response inhibitor. The peptides are useful for
CC	treating <i>S. aureus</i> infections.
XX	
SO	Sequence 9 AA;
QY	Query Match 82.1%; Score 32; DB 23; Length 9;
	Best Local Similarity 88.9%; Pred. No. 9.3e+05;
	Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
DB	1 GVNAXSSAF 9
	1 GVNAXSLP 9
XX	
AC	AAW50899;
XX	
DT	08-MAY-2002 (first entry)
XX	
DE	AgrD-autoinducing cyclic peptide, inhibitor of agr response.
XX	
KM	<i>Staphylococcus aureus</i> ; AgrD; agr response; inhibitor; antibiotic;
XX	antibacterial; infection; therapy; cyclic.
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 5 /note= "any amino acid"
FT	Misc-difference 5 /note= "note linked to residue 9 to form cyclic
FT	peptide"
FT	Misc-difference 9 /note= "note linked to residue 5 to form cyclic
FT	peptide"
XX	
PN	US6337385-B1.
XX	
PD	08-JAN-2002.
XX	
PF	24-JUN-1999; 99US-0339511.
XX	
PR	24-JUN-1998; 98US-090402P.
XX	
PA	(UYRO) UNITV ROCKEFELLER.
XX	(UYNY) UNITV NEW YORK STATE.
PI	
XX	
DR	Mult TM, Mayville P, Novick RP, Beavis R, Ji G;
XX	
XX	WPI, 2002-170774/22.
XX	
PT	Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
PT	interference and for treating <i>Staphylococcus aureus</i> infection in a
PT	subject -
XX	
PS	Claim 7; Column 19; 18pp; English.
XX	
CC	The present sequence is that of a novel synthetic cyclic peptide
CC	of the invention that is capable of inhibiting the agr response of
CC	<i>Staphylococcus aureus</i> . It is an AgrD-autoinducing peptide, where

CC AgrD is a secreted agr-encoded peptide and where the agr locus
 CC controls the synthesis of virulence factor and other extracellular
 CC proteins responsible for pathogenicity in *S. aureus*. Preferred
 CC peptides may have the sequence NH₂-X(n)-Z-X(Y)-COOH, with a
 CC cyclic bond between the Z residue and COOH other than a thioester
 CC bond, where X is an amino acid, an amino acid analogue, a
 CC peptidomimetic or non-amide isostere, Z is a synthetic or a
 CC biosynthetic amino acid, n is 0-10 and Y is 1-10. The cyclic bond
 CC is especially a lactam or lactone bond. The thiololactone
 CC structure within native AgrD peptides is required for activation
 CC of the agr response. Elimination of the thiol ester component of
 CC the cyclic ring structure can destroy agr response activating
 CC activity while preserving and enhancing inhibitory activity. A
 CC claimed method of preparing a cyclic peptide involves: assembling
 CC a linear peptide chain on to a solid phase resin support;
 CC protecting the resulting protected assembled peptide; treating the
 CC deprotected peptide with neutral buffer for a time sufficient to
 CC form the cyclic peptide and cleave the peptide from the support;
 CC and recovering the cyclic peptide. The peptide is useful for
 CC bacterial interference, especially for the treatment of *S. aureus*
 CC infection.

CC Sequence 9 AA;

Query Match 82.1%; Score 32; DB 23; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
 1 GVNAXSSLP 9

DB

RESULT 15

AAMS0908 AAMS0908 standard; Peptide; 9 AA.

AC AAMS0908;

DT 08-MAY-2002 (first entry)

XX Protected peptide used in cyclic peptide production.

XX Staphylococcus aureus; AgrD; agr response; inhibitor; antibiotic;

KW antibacterial; infection; therapy; cyclic.

OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "Z-Gly"

FT Modified-site 5 /label= Dpr(Boc)

FT Modified-site 6 /note= "Ser(Bz1)"

FT Modified-site 7 /note= "Ser(Bz1)"

FT Misc-difference 5 /note= "note linked to residue 9 to form cyclic peptide"

FT Misc-difference 9 /note= "note linked to residue 5 to form cyclic peptide"

FT

PN US6337385-B1.

PD 06-JAN-2002.

PF 24-JUN-1999; 99US-0339511.

PR 24-JUN-1998; 98US-090402P.

PA (UVRQ) UNIV ROCKEFELLER.

PA (UVRQ) UNIV NEW YORK STATE.
 XX
 XX Mair TW, Mayville P, Novick RP, Beavis R, Ji G;
 XX
 XX WPI; 2002-170774/22.

PT Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
 PT interference and for treating Staphylococcus aureus infection in a
 PT subject -
 XX
 XX Disclosure; Column 14; 18pp; English.

CC The present sequence is that of a protected peptide used in an
 CC example of the preparation of novel synthetic cyclic peptides of
 CC the invention (see AAMS0899-908). The peptide corresponds to the
 CC Staphylococcus aureus AgrDI sequence with a Cys5 to diaminopropionic
 CC acid (Dpr) mutation (lactam). It was synthesised on a Wang-resin
 CC using an Fmoc N-alpha protection strategy. Following chain assembly,
 CC the peptide was cleaved from the support and the Dpr-5 residue
 CC deprotected by treatment with a trifluoroacetic acid:anisole:water
 CC mixture (90:5:5) for 4 hr. The partially protected peptide-alpha
 CC carboxylates were then dissolved in DMF and treated with PyBOP.
 CC Cyclization was complete after 2 hr. The remaining protecting groups
 CC were removed by treatment with HF and the peptide purified by HPLC.
 CC The cyclic peptide is capable of inhibiting the agr response of
 CC Staphylococcus aureus. The thiololactone structure within native
 CC AgrD peptides is required for activation of this response.
 CC Replacement of the thiol ester component of the cyclic ring
 CC structure with a lactam (as in the present case) or a lactone can
 CC destroy agr response activating activity while preserving and
 CC enhancing inhibitory activity. The cyclic peptides are useful for
 CC bacterial interference, especially for the treatment of *S. aureus*
 CC infection.

SQ Sequence 9 AA;

Query Match 82.1%; Score 32; DB 23; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
 1 GVNAXSSLP 9

DB

Search completed: January 29, 2004, 14:52:13
 Job time : 35.0735 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:46:49 ; Search time 11.1176 Seconds
(without alignments)

77.851 Million cell updates/sec

Title: US-10-032-950-7
Perfect score: 39
Sequence: 1 GVNAXSSAF 9

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	31	79.5	47	2	C89995	AgRD protein [impo
2	31	79.5	924	2	B95192	hypothetical prote
3	31	79.5	924	2	G98058	hypothetical prote
4	30	76.9	163	2	A87066	conserved hypochet
5	30	76.9	167	2	C70866	hypothetical prote
6	30	76.9	169	2	T45243	hypothetical prote
7	30	76.9	182	2	A28067	lysosomal membrane
8	30	76.9	405	2	A60534	P2B/LAMP-1 precursor
9	30	76.9	407	2	A30200	120K lysosomal mem
10	30	76.9	416	1	A31959	lysosome-associated
11	30	76.9	500	1	EPFF	zip protein precursor
12	30	76.9	529	2	T23740	hypothetical prote
13	29	74.4	225	2	T50707	uracase accessory p
14	29	74.4	270	1	A38737	gap junction prote
15	29	74.4	270	2	G23590	connexin 31 - mous
16	29	74.4	270	2	UE0274	connexin 31 - huma
17	29	74.4	271	1	B43433	gap junction prote
18	29	74.4	271	1	D42053	gap junction prote
19	29	74.4	316	2	B69315	acetylcholinestere
20	29	74.4	333	1	G69812	ferriochrome ABC tr
21	29	74.4	411	2	S75327	lysostaphin - syne
22	29	74.4	431	2	AD0337	hypothetical prote
23	29	74.4	431	2	G98848	hypothetical prote
24	29	74.4	653	2	T03319	gene 112 protein -
25	29	74.4	3635	2	T10053	lamtin alpha 5 ch
26	28	71.8	88	2	A82434	conserved hypotnet
27	28	71.8	114	2	S50624	hypothetical prote
28	28	71.8	189	2	AC3420	phage minor tail p
29	28	71.8	206	2	A65189	yfgW protein - Esc

30	28	71.8	260	2	A91225	cytoplasmic Diase
31	28	71.8	260	2	AF0915	probable deoxyribo
32	28	71.8	264	2	G86071	hypothetical prote
33	28	71.8	326	2	B86832	oxidoreductase yrb
34	28	71.8	332	2	T04754	hypothetical prote
35	28	71.8	346	2	AC0187	probable iron-side
36	28	71.8	364	2	S43574	C05B5.3 protein (c
37	28	71.8	398	2	F85056	hypothetical prote
38	28	71.8	401	2	C88571	protein C05B5.3 (l
39	28	71.8	453	2	S18597	tubulin beta chain
40	28	71.8	466	2	G96713	probable DNA-bind
41	28	71.8	493	1	S51574	mocr protein - Rhi
42	28	71.8	508	2	A32225	nerve growth facto
43	28	71.8	512	2	H70662	probable pica prot
44	28	71.8	523	2	G86349	hypothetical prote
45	28	71.8	540	2	S35071	auxin-resistance p

ALIGNMENTS

RESULT 1
C89995
AgRD protein [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C/Accession: C89995
R/Rutoda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Chii, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsutsu, K.
Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89758; MUID:21311952; PMID:11418146
A/Accession: C89995
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-47 <KUR>
A/Cross-references: GB:BA00016; PIR:G13701631; PIRN:BA83124.1; GSPDB:GN00149
A/Experimental source: strain N315
C/Genetics:
A/Gene: agrD

Query Match
Best Local Similarity 79.5%; Score 31; DB 2; Length 47;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
DB 24 GVNAXSSAF 32

RESULT 2
B95192
hypothetical protein SPI652 (imported) - Streptococcus pneumoniae (strain TIGR4)
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C/Accession: B95192
R/Rutella, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
nson, J.D.; Umayan, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple
nson, T.; Hickey, E.K.; Holt, I.E.
Science 283, 498-506, 2001
A/Authors: Loftus, B.U.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A/Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A/Reference number: A95000; MUID:21357209; PMID:11463516
A/Accession: B95192
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-924 <KUR>
A/Cross-references: GB:AE005672; PIRN:PAK75731.1; PIR:G14973142; GSPDB:GN00164; TIGR:
A/Experimental source: strain TIGR4
C/Genetics:
A/Gene: SPI652

Query Match 79.5%; Score 31; DB 2; Length 924;
 Best Local Similarity 66.7%; Pred. No. 45;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAKSSAF 9
 DB 239 GLNAFSSAY 247

RESULT 3
 G98058
 hypothetical protein spr1496 [imported] - Streptococcus pneumoniae (strain R6)
 C/Species: Streptococcus pneumoniae
 C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
 C/Accession: G98058
 R/Authors: J.A.; Albom Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A/Title: Genome of the Bacterium Streptococcus pneumoniae strain R6.
 A/Reference number: A97872; MUID:21429245; PMID:11544234
 A/Accession: G98058
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-924 <KUR>
 A/Cross-references: GB:AEO07317; PIDN:AAL00300.1; PID:G15459156; GSPDB:GN00174
 C/Genetics:
 A/Gene: spr1496

Query Match 79.5%; Score 31; DB 2; Length 924;
 Best Local Similarity 66.7%; Pred. No. 45;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAKSSAF 9
 DB 239 GLNAFSSAY 247

RESULT 4
 A87066
 conserved hypothetical protein ML1255 [imported] - Mycobacterium leprae
 C/Species: Mycobacterium leprae
 C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C/Accession: A87066
 R/Cole, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
 R.; Davies, R.M.; Devlin, K.; Duthey, S.; Felkwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
 eam, M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
 A/Title: Massive gene decay in the leprosy bacillus.
 A/Reference number: A86909; MUID:21128732; PMID:11234002
 A/Accession: A87066
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-163 <STO>
 A/Cross-references: GB:AL450380; NID:G13093193; PIDN:CAC31636.1; GSPDB:GN00147
 C/Genetics:
 A/Gene: ML1255
 C/Superfamily: Mycobacterium leprae hypothetical protein MLCB1610.16

Query Match 76.9%; Score 30; DB 2; Length 163;
 Best Local Similarity 77.8%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAKSSAF 9
 DB 132 GVAAASSAF 140

RESULT 5
 C70866
 hypothetical protein Rv2468c - Mycobacterium tuberculosis (strain H37Rv)

C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 28-Jul-2000
 C/Accession: C70866
 R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
 , Connor, R.; Davies, R.; Devlin, K.; Felkwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
 A/Reference number: A70500; MUID:98295987; PMID:9634230
 A/Accession: C70866
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-167 <COL>
 A/Cross-references: GB:AL021246; GB:AL123456; NID:G3261507; PIDN:CAA16045.1; PID:G279
 A/Experimental source: strain H37Rv
 C/Genetics:
 A/Gene: Rv2468c
 C/Superfamily: Mycobacterium leprae hypothetical protein MLCB1610.16

Query Match 76.9%; Score 30; DB 2; Length 167;
 Best Local Similarity 77.8%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAKSSAF 9
 DB 136 GVAAASSAF 144

RESULT 6
 T45243
 hypothetical protein MLCB1610.16 [imported] - Mycobacterium leprae
 C/Species: Mycobacterium leprae
 C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 28-Jul-2000
 C/Accession: T45243
 R/James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, March 1999
 A/Reference number: Z22949
 A/Accession: T45243
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-169 <JAM>
 A/Cross-references: EMBL:AL049913; PIDN:CAB43162.1
 A/Experimental source: cosmid B1610
 C/Genetics:
 A/Gene: MLCB1610.16
 C/Superfamily: Mycobacterium leprae hypothetical protein MLCB1610.16

Query Match 76.9%; Score 30; DB 2; Length 169;
 Best Local Similarity 77.8%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAKSSAF 9
 DB 138 GVAAASSAF 146

RESULT 7
 A28067
 lysosomal membrane glycoprotein LAMP-1 - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 20-Aug-1999
 C/Accession: A28067
 R/Chen, J.W.; Cha, Y.; Yukse, K.U.; Gracy, R.W.; August, J.T.
 J. Biol. Chem. 263, 8754-8758, 1988
 A/Title: Isolation and sequencing of a cDNA clone encoding lysosomal membrane glycopro
 A/Reference number: A28067; MUID:88243732; PMID:3379044
 A/Accession: A28067
 A/Molecule type: mRNA
 A/Residues: 1-362 <CHE>
 A/Cross-references: GB:J03881; NID:G198706; PIDN:AAA39411.1; PID:G293692
 A/Note: the authors translated the codon ATT for residue 1 as Iau and CCG for residue
 C/Superfamily: lysosome-associated membrane protein

C/Keywords: glycoprotein; membrane protein

Query Match 76.9%; Score 30; DB 2; Length 382;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 2; Gaps 0;

QY 1 GVNAXSSAF 9
DB 256 GVNAXSSSLF 264

RESULT 8

P2B/LAMP-1 precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 03-May-1996
C/Accession: A60534
R/Hefner, M.; Yousefi, S.; Dennis, J.W.
Cancer Res. 49, 6077-6084, 1989
A/Title: Molecular characterization of P2B/LAMP-1, a major protein target of a metastatic
A/Reference number: A60534; MUID:90002989; PMID:2676155
A/Accession: A60534
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-405 <HEF>
C/Superfamily: lysosome-associated membrane protein

Query Match 76.9%; Score 30; DB 2; Length 405;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 2; Gaps 0;

QY 1 GVNAXSSAF 9
DB 279 GVNAXSSSLF 287

RESULT 9

A30200
120K lysosomal membrane glycoprotein precursor - rat
N/Alternate names: staloglycoprotein
C/Species: Rattus norvegicus (Norway rat)
C/Date: 18-Oct-1989 #sequence_revision 30-Sep-1991 #text_change 20-Aug-1995
C/Accession: A30200; S03331
R/Howe, C.L.; Granger, B.L.; Hull, M.; Green, S.A.; Gabel, C.A.; Helenius, A.; Mellman, P.
Proc. Natl. Acad. Sci. U.S.A. 85, 7577-7581, 1988
A/Title: Derived protein sequence, oligosaccharides, and membrane insertion of the 120-k
glycoproteins.
A/Reference number: A30200; MUID:89017240; PMID:3174652
A/Accession: A30200
A/Molecule type: mRNA
A/Residues: 1-407 <HOW>
A/Cross-references: EMBL:J03672
A/Note: the authors translated the codon GGG for residue 15 as Val
R/Himeno, M.; Noguchi, Y.; Sasaki, H.; Tanaka, Y.; Furuno, K.; Kono, A.; Sakaki, Y.; Kat
EBS Lett. 244, 351-356, 1989
A/Title: Isolation and sequencing of a cDNA clone encoding 107 kDa staloglycoprotein in
A/Reference number: S03331; MUID:89153580; PMID:2920835
A/Accession: S03331
A/Molecule type: mRNA
A/Residues: 22-407 <HIM>
A/Cross-references: EMBL:X14765; NID:G56577; PIDN:CA32873.1; PID:G56578
A/Note: part of this sequence, including the amino end of the mature protein, was confir
C/Superfamily: lysosome-associated membrane protein
C/Keywords: glycoprotein; membrane protein
F/22-407/Product: lysosomal membrane glycoprotein, 107K #status experimental <MAT>

Query Match 76.9%; Score 30; DB 2; Length 407;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 2; Gaps 0;

QY 1 GVNAXSSAF 9
DB 281 GVNAXSSSLF 289

RESULT 10

A31959
Lysosome-associated membrane protein 1 precursor - human
N/Alternate names: lysosomal-associated membrane protein lamp A; major lysosomal mem
C/Species: Homo sapiens (man)
C/Date: 31-Mar-1990 #sequence_revision 01-Dec-1995 #text_change 22-Jun-1999
C/Accession: A31959; B46044; S02150; A23656; A32685; S34737; A30210; B30210
R/Fukuda, M.; Vittal, J.; Matteson, J.; Carlsson, S.R.
J. Biol. Chem. 263, 18920-18928, 1988
A/Title: Cloning of cDNAs encoding human lysosomal membrane glycoproteins, h-lamp-1 ar
A/Reference number: A32659; MUID:89066687; PMID:3198605
A/Accession: A31959
A/Molecule type: mRNA
A/Residues: 1-416 <FUK>
A/Cross-references: GB:J04182; NID:G186927; PIDN:AAA60382.1; PID:G307109
R/Sawada, R.; Jardine, K.A.; Fukuda, M.
J. Biol. Chem. 268, 9014-9022, 1993
A/Title: The genes of major lysosomal membrane glycoproteins, lamp-1 and lamp-2, 5'-f
A/Reference number: A46044; MUID:93232065; PMID:8517882
A/Accession: B46044
A/Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A/Molecule type: DNA
A/Residues: 1-416 <SAW>
A/Note: sequence extracted from NCBI backbone (NCBI:P:129678)
R/Mane, S.M.; Marzella, L.; Baiton, D.F.; Holt, V.K.; Cha, Y.; Hildreth, J.E.K.; Aug
Arch. Biochem. Biophys. 268, 360-378, 1989
A/Title: Purification and characterization of human lysosomal membrane glycoproteins.
A/Reference number: S02150; MUID:89104438; PMID:2912382
A/Accession: S02150
A/Molecule type: protein
A/Residues: 28-35, 'X', 37-39, 'X', 41-43, 'X', 45-58 <MAN>
R/Carlsson, S.R.; Fukuda, M.
J. Biol. Chem. 265, 20488-20495, 1990
A/Title: The polyclacotaminoglycans of human lysosomal membrane glycoproteins lamp-1
A/Reference number: A23656; MUID:91056099; PMID:2243102
A/Accession: A23656
A/Status: preliminary
A/Molecule type: protein
A/Residues: 57-71, 116-135, 215-234 <CAR>
R/Carlsson, S.R.; Fukuda, M.
J. Biol. Chem. 264, 20526-20531, 1989
A/Title: Structure of human lysosomal membrane glycoprotein 1. Assignment of disulfid
A/Reference number: A32685; MUID:90062189; PMID:2584229
A/Accession: A32685
A/Molecule type: protein
A/Residues: 31-35, 72-74, 'X', 76-77, 144-153, 177-179, 'X', 181-187, 215-219, 255-259, 'X', 261
R/Carlsson, S.R.; Lyckxell, P.O.; Fukuda, M.
Arch. Biochem. Biophys. 304, 65-73, 1993
A/Title: Assignment of O-glycan attachment sites to the hinge-like regions of human 1
A/Reference number: S34737; MUID:93312023; PMID:8323299
A/Accession: S34737
A/Molecule type: protein
A/Residues: 190-214 <C3>
R/Vittal, J.; Carlsson, S.R.; Siebert, P.D.; Fukuda, M.
Proc. Natl. Acad. Sci. U.S.A. 85, 3743-3747, 1988
A/Title: Molecular cloning of cDNAs encoding lamp A, a human lysosomal membrane glyco
A/Reference number: A30210; MUID:88234502; PMID:3131762
A/Accession: A30210
A/Molecule type: mRNA
A/Residues: 40-416 <VIT>
A/Cross-references: GB:J03263; NID:G187178; PIDN:AAA59524.1; PID:G307132
A/Accession: B30210
A/Molecule type: protein
A/Residues: 116-119, 'X', 121-128, 'X', 130-131, 'X', 133-135, 163, 'X', 165-179, 'X', 181-189, '
C/Genetics:
A/Gene: GDB:LAMP1
A/Cross-references: GDB:120137; OMIM:153330
A/Map position: 13q34-13q34
C/Superfamily: lysosome-associated membrane protein
C/Keywords: glycoprotein; lysosome; transmembrane protein

F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-416/Product: lysosomal membrane glycoprotein 1 #status predicted <WAT>
 F:28-381/Domain: lysosomal lumenal #status predicted <LVS>
 F:389-405/Domain: transmembrane #status predicted <TM>
 F:406-416/Domain: intracellular #status predicted <CIT>
 F:46-44,83,102,106,240,248,292,321,379/Binding site: carbohydrate (asn) (covalent) #stat
 F:40-79,154-190,230-268,337-374/Disulfide bonds: #status experimental
 F:61,75,120,129,164,180,222,227,260/Binding site: carbohydrate (asn) (covalent) #status
 F:196/Binding site: carbohydrate (Ser) (covalent) (partial) #status experimental
 F:198,199/Binding site: carbohydrate (Thr) (covalent) #status experimental
 F:206,208,210/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 76.9%; Score 30; DB 1; Length 416;
 Best Local Similarity 66.7%; Pred. No. 33;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Cy 1 GVNAXSSAF 9
 Db 290 GVNAXSSRF 298

RESULT 11
 EPRF
 zip protein precursor - fruit fly (Drosophila melanogaster)
 N/Alternate names: zipper protein
 C/Species: Drosophila melanogaster
 C/Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 22-Jun-1999
 C/Accession: S00483; A37532
 R/Zhao, D.B.; Cole, S.; Jaehnis, F.; Haller, J.; Jaekle, H.
 EMBL J. 7, 1115-1119, 1988
 A/Title: Zipper encodes a putative integral membrane protein required for normal axon pa
 A/Reference number: S00483; MUID:86226414; PMID:3402433
 A/Accession: S00483
 A/Molecule type: DNA
 A/Residues: 1-500 <ZHA>
 A/Cross-references: GB:X07450; NID:98854; PIDN:CAA30332.1; PID:98855
 A/Accession: A37532
 A/Molecule type: mRNA
 A/Residues: 1-500 <ZHA1>
 A/Cross-references: EMBL:X07450; NID:98854; PIDN:CAA30332.1; PID:98855
 C/Genetics:
 A/Gene: zip
 A/Cross-references: FlyBase:FBgn0004055
 A/Introns: 237/1; 354/1
 C/Superfamily: zip protein
 C/Keywords: glycoprotein, transmembrane protein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-500/Product: zip protein #status predicted <WAT>
 F:22-470/Domain: extracellular #status predicted <EXT>
 F:466-486/Domain: transmembrane #status predicted <TM>
 F:496-500/Domain: intracellular #status predicted <INT>
 F:35,237,317,374,448/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 76.9%; Score 30; DB 1; Length 500;
 Best Local Similarity 66.7%; Pred. No. 40;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy 1 GVNAXSSAF 9
 Db 463 GVNAXGSTF 471

RESULT 12
 T23740
 hypothetical protein M106.3 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T23740
 R/Palmer, S.
 submitted to the EMBL Data Library, December 1994
 A/Reference number: Z19792
 A/Accession: T23740
 A/Status: preliminary; translated from GB/EWBL/DBJ

A/Molecule type: DNA
 A/Residues: 1-529 <M1>
 A/Cross-references: EMBL:Z46935; PIDN:CAA87050.1; GSPDB:GN00020; CESP:M106.3
 A/Experimental source: clone M106
 C/Genetics:
 A/Gene: CESP:M106.3
 A/Map position: 2
 A/Introns: 34/1; 175/3; 224/3; 271/3; 341/3; 372/1; 401/3; 446/3; 475/3; 494/3

Query Match 76.9%; Score 30; DB 2; Length 529;
 Best Local Similarity 66.7%; Pred. No. 43;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Cy 1 GVNAXSSAF 9
 Db 349 GVNAXSAF 357

RESULT 13
 T50707
 urease accessory protein D [imported] - Rhodobacter sphaeroides
 C/Species: Rhodobacter sphaeroides
 C/Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C/Accession: T50707
 R/Choudhary, M.; Kaplan, S.
 Nucleic Acids Res. 28, 862-867, 2000
 A/Title: DNA sequence analysis of the photosynthesis region of Rhodobacter sphaeroides
 A/Reference number: Z25222; MUID:20115911; PMID:10648776
 A/Accession: T50707
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-225 <CHO>
 A/Cross-references: EMBL:AF195122; PIDN:AAF24251.1
 A/Experimental source: strain 2.4.1
 C/Genetics:
 A/Gene: ureD

Query Match 74.4%; Score 29; DB 2; Length 225;
 Best Local Similarity 66.7%; Pred. No. 30;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Cy 1 GVNAXSSAF 9
 Db 179 GVNAXSAF 187

RESULT 14
 A38737
 gap junction protein Cx31 - rat
 N/Alternate names: connexin-31
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 13-Sep-1991 #sequence_revision 02-Jun-1994 #text_change 22-Jun-1999
 C/Accession: A38737
 R/Hoh, J.H.; John, S.A.; Revel, J.P.
 J. Biol. Chem. 266, 6524-6531, 1991
 A/Title: Molecular cloning and characterization of a new member of the gap junction g
 A/Reference number: A38737; MUID:9117912; PMID:1706719
 A/Accession: A38737
 A/Molecule type: DNA
 A/Residues: 1-270 <CHO>
 A/Cross-references: GB:M59936; NID:9203659; PIDN:AAA40997.1; PID:9203660
 A/Note: the authors translated the codon CAG for residue 7 as Gly and GAC for residue
 C/Superfamily: gap junction protein
 C/Keywords: gap junction; phosphoprotein; placenta; transmembrane protein
 F:1-20/Domain: intracellular #status predicted <INT>
 F:21-40/Domain: transmembrane #status predicted <TM>
 F:41-73/Domain: extracellular #status predicted <EXT>
 F:74-96/Domain: transmembrane #status predicted <TM>
 F:97-122/Domain: intracellular #status predicted <INT>
 F:123-159/Domain: transmembrane #status predicted <TM>
 F:160-184/Domain: extracellular #status predicted <EXT>
 F:185-207/Domain: transmembrane #status predicted <TM>
 F:208-270/Domain: intracellular #status predicted <INT>

Query Match 74.4%; Score 29; DB 1; Length 270;
 Best Local Similarity 66.7%; Pred. No. 36;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYNAXSAF 9
 DB 12 GVNQYSTAF 20

RESULT 15

S23590
 comexin 31 - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 13-Aug-1999
 C/Accession: S23590
 R/Hennemann, H.; Schwarz, H.J.; Willecke, K.
 Eur. J. Cell Biol. 57, 51-58, 1992
 A/Title: Characterization of gap junction genes expressed in F9 embryonic carcinoma cell
 A/Reference number: S23589; MUID:92347372; PMID:1322300
 A/Accession: S23590
 A/Molecule type: mRNA
 A/Residues: 1-270 <HEN>
 A/Cross-references: EMBL:X63099; NID:G50598; PIDN:CAA44813.1; P/D:G50599
 C/Superfamily: gap junction protein
 C/Keywords: transmembrane protein

Query Match 74.4%; Score 29; DB 2; Length 270;
 Best Local Similarity 66.7%; Pred. No. 36;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYNAXSAF 9
 DB 12 GVNQYSTAF 20

Search completed: January 29, 2004, 14:58:24
 Job time : 12.1176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:45:44 ; Search time 5.82353 Seconds

(without alignments)
72.678 Million cell updates/sec

Title: US-10-032-950-7

Perfect score: 39

Sequence: 1 GVNXXSSAF 9

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	76.9	406	1	LMP1_MOUSE
2	30	76.9	407	1	LMP1_RAT
3	30	76.9	416	1	LMP1_HUMAN
4	30	76.9	488	1	UZ1P_DROME
5	29	74.4	270	1	CBX3_HUMAN
6	29	74.4	270	1	CBX3_MOUSE
7	29	74.4	270	1	CBX3_RAT
8	29	74.4	271	1	CBX5_MOUSE
9	29	74.4	271	1	CBX5_RAT
10	29	74.4	273	1	CBX5_HUMAN
11	29	74.4	3718	1	LMAS_MOUSE
12	28	71.8	114	1	YEY1_YEAST
13	28	71.8	260	1	TATD_ECOLI
14	28	71.8	401	1	YK03_CAEEL
15	28	71.8	493	1	MOCR_RHIME
16	28	71.8	508	1	EGRI_RAT
17	28	71.8	520	1	PHLA_MYCTU
18	28	71.8	532	1	CBR1_DROAC
19	28	71.8	533	1	EGRI_MOUSE
20	28	71.8	540	1	AXR1_ARATH
21	28	71.8	545	1	NUSM_ALBEO
22	28	71.8	646	1	GYRB_MYCGA
23	28	71.8	650	1	GYRB_MYCGE
24	28	71.8	650	1	GYRB_MYCPN
25	28	71.8	694	1	PPK_CAMEJ
26	28	71.8	857	1	MEB3_GALA
27	27	69.2	180	1	MP17_LYMSI
28	27	69.2	222	1	YS40_MYCPU
29	27	69.2	237	1	YMR9_YEAST
30	27	69.2	269	1	PROC_ECOLI
31	27	69.2	287	1	Y552_HELPY
32	27	69.2	289	1	Y552_HELPY
33	27	69.2	317	1	UL14_HEVERB

34	27	69.2	330	1	GPDA_THETN
35	27	69.2	339	1	PE29_ARATH
36	27	69.2	359	1	ADD2_STRCO
37	27	69.2	365	1	ADD_MYCTU
38	27	69.2	382	1	MTLD_ECOL6
39	27	69.2	382	1	MTLD_ECOLI
40	27	69.2	382	1	MTLD_KUBPN
41	27	69.2	382	1	MTLD_SALTY
42	27	69.2	382	1	MTLD_SALTY
43	27	69.2	388	1	SCB3_SCHPO
44	27	69.2	389	1	KSS5_ECOLI
45	27	69.2	396	1	ADD5_STRCO

ALIGNMENTS

RESULT 1
LMP1_MOUSE STANDARD; PRT; 406 AA.
ID LMP1_MOUSE
AC P1438; Q62020; (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 15, Last sequence update)
DT 01-AUG-1990 (Rel. 35, Last annotation update)
DE Lysosome-associated membrane glycoprotein 1 precursor (LAMP-1) (LGP-A)
DE (LGP-120) (CD107A) (P28).
GN LAMP1 OR LAMP-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90307736; PubMed=2142158;
RA Granger B.L., Green S.A., Gabel C.A., Howe C.L., Mellman I.,
RA Helenius A.;
RT "Characterization and cloning of lgp110, a lysosomal membrane
RT glycoprotein from mouse and rat cells.";
RL J. Biol. Chem. 265:12036-12043(1990).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88243732; PubMed=3379044;
RA Chen J.W., Cha Y., Yukel K.U., Gracy R.W., August J.T.;
RT "Isolation and sequencing of a cDNA clone encoding lysosomal membrane
RT glycoprotein mouse LAMP-1. Sequence similarity to proteins bearing
RT onco-differentiation antigens.";
RL J. Biol. Chem. 263:8754-8758(1988).
[4]
RP DISULFIDE BONDS.
RX MEDLINE=90237040; PubMed=2332434;
RA Artshorn L.M., Barles B.J., August J.T.;
RT "The disulfide structure of mouse lysosome-associated membrane
RT protein 1.";
RL J. Biol. Chem. 265:7419-7423(1990).
[1]
CC FUNCTION: PRESENTS CARBOHYDRATE LIGANDS TO SELECTINS. ALSO
CC IMPLICATED IN TUMOR CELL METASTASIS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Lysosomal.
CC -1- THIS PROTEIN SHUTTLES BETWEEN LYSOSOMES, ENDOSOMES, AND THE PLASMA
CC MEMBRANE.
CC -1- PTM: O- AND N-GLYCOSYLATED; SOME OF THE N-GLYCANS ATTACHED TO
CC LAMP-1 ARE POLYLACTOSAMINOGLYCANS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LAMP FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL, subscription
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

FT	CABOHYD	215	215	N-LINKED (GLCNAC .)	(POTENTIAL) .
FT	CABOHYD	220	220	N-LINKED (GLCNAC .)	(POTENTIAL) .
FT	CABOHYD	233	233	N-LINKED (GLCNAC .)	(POTENTIAL) .
FT	CABOHYD	241	241	N-LINKED (GLCNAC .)	(POTENTIAL) .
FT	CABOHYD	271	271	N-LINKED (GLCNAC .)	(POTENTIAL) .
FT	CABOHYD	283	283	N-LINKED (GLCNAC .)	(POTENTIAL) .
FT	CABOHYD	297	297	N-LINKED (GLCNAC .)	(POTENTIAL) .
FT	CABOHYD	312	312	N-LINKED (GLCNAC .)	(POTENTIAL) .
FT	CONFLICT	293	294	GV -> EF (IN REF. 3).	
FT	CONFLICT	329	329	N -> T (IN REF. 3).	
FT	CONFLICT	356	357	SD -> VT (IN REF. 3).	
SQ	SEQUENCE	407 AA;	43969 NM;	2594790749A7C88 CRC64;	
QY	Query Match	Best Local Similarity	76.9%;	Score 30;	DB 1;
	Matches	6;	Conservative	1;	Mismatches
				2;	Indels
					Gaps
Db	281 GNMATSLF	289			
RESULT 3					
LMP1_HUMAN					
ID	LMP1_HUMAN	STANDARD;	PRT;	416 AA.	
AC	P11279;				
DT	01-JUL-1989 (Rel. 11, Created)				
DT	01-JAN-1990 (Rel. 13, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Lysosome-associated membrane glycoprotein 1 precursor (LAMP-1)				
DE	(CD107a antigen).				
GN	LAMP1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RX	SEQUENCE FROM N.A., AND SEQUENCE OF 26-50; 116-130 AND 163-197.				
RA	Fukuda M.; Vitalita J.; Matteson J.; Carlsson S.R.;				
RT	"Cloning of cDNAs encoding human lysosomal membrane glycoproteins, h-				
RT	lamp-1 and h-lamp-2. Comparison of their deduced amino acid				
RT	sequences."				
RL	J. Biol. Chem. 263:18920-18928(1988).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Placenta;				
RX	MEDLINE=93232065; PubMed=8517882;				
RA	Sawada R.; Jardine K.A.; Fukuda M.;				
RT	"The genes of major lysosomal membrane glycoproteins, lamp-1 and				
RT	lamp-2, 5'-flanking sequence of lamp-2 gene and comparison of exon				
RT	organization in two genes."				
RL	J. Biol. Chem. 268:9014-9022(1993).				
RN	[3]				
RP	SEQUENCE OF 32-416 FROM N.A., AND PARTIAL SEQUENCE.				
RX	MEDLINE=88234502; PubMed=311752;				
RA	Vitalita J.; Carlsson S.R.; Siebert P.D.; Fukuda M.;				
RT	"Molecular cloning of cDNAs encoding lamp A, a human lysosomal				
RT	membrane glycoprotein with apparent Mr approximately equal to				
RT	lipo. 120,000."				
RL	Proc. Natl. Acad. Sci. U.S.A. 85:3743-3747(1988).				
RN	[4]				
RP	SEQUENCE OF 28-58.				
RX	MEDLINE=89104438; PubMed=2912382;				
RA	Mane S.M.; Marzella L.; Bainton D.F.; Holt V.K.; Cha Y.;				
RA	Hildreth J.E.K.; August J.T.;				
RT	"Purification and characterization of human lysosomal membrane				
RT	glycoproteins."				
RL	Arch. Biochem. Biophys. 268:360-378(1989).				
RN	[5]				
RP	DISTUPTDE BONDS.				
RX	MEDLINE=90062189; PubMed=2564225;				
RA	Carlsson S.R.; Fukuda M.;				

RP			"structure of human lysosomal membrane glycoprotein 1. Assignment of
RT			disulfide bonds and visualization of its domain arrangement.";
RL	J. Biol. Chem.	264:20526-20531(1989).	
RN			POLYLACTOSAMINOGLYCANS.
RA	MEDLINE=91056099; PubMed=2243102;		
RX	Carlsson S.R., Fukuda M.;		
RT	"Assignment of O-glycan attachment sites to the hinge-like regions of		
RL	human lysosomal membrane glycoproteins lamp-1 and lamp-2.";		
Arch. Biochem. Biophys.	304:65-77(1993).		
CC	-I- FUNCTION: PRESENTS CARBOHYDRATE LIGANDS TO SELECTINS. ALSO		
CC	IMPLICATED IN TUMOR CELL METASTASIS.		
CC	-I- SUBCELLULAR LOCATION: Type I membrane protein, lysosome1.		
CC	THIS PROTEIN SHUTTLES BETWEEN LYSSOMES, ENDOSOMES, AND THE PLASMA		
CC	MEMBRANE.		
CC	-I- PTM: O- AND N-GLYCOSYLATED; SOME OF THE 18 N-GLYCANS ATTACHED TO		
CC	LAMP-1 ARE POLYLACTOSAMINOGLYCANS.		
CC	-I- SIMILARITY: BELONGS TO THE LAMP FAMILY.		
CC	-I- DATABASE: NAME=PROW; NOTE=CD guide CD107a entry/;		
CC	WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd107a.htm".		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by, and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; J04183; AAA60382.1; -		
DR	EMBL; J03263; AAA59524.1; -		
DR	PIR; A31959; A31959.		
DR	GlycoSuiteDB; P11793; -		
DR	Genew; HGNC:6499; LAMP1.		
DR	MIM; 153330; -		
DR	GO; GO:0005887; C:integral to plasma membrane; TAS.		
DR	GO; GO:0005764; C:lysosome; TAS.		
DR	GO; GO:0005624; C:membrane fraction; TAS.		
DR	InterPro; IPRO02000; Lamp.		
DR	Pfam; Pf01299; Lamp; 1.		
DR	PRINTS; PR00336; LYSASSOCTDMP.		
DR	PROSITE; PS00310; LAMP 1; 2.		
KW	PROSITE; PS00311; LAMP 2; 1.		
Transmembrane; Glycoprotein; Lysosome; Signal.			
FT	SIGNAL	1..27	
FT	CMAIN	28..416	LYSSOME-ASSOCIATED MEMBRANE GLYCOPROTEIN
FT			1.
FT	DOMAIN	28..381	LUMENAL (POTENTIAL).
FT	TRANSYM	382..404	POTENTIAL.
FT	DOMAIN	405..416	CYTOSOLSMIC (POTENTIAL).
FT	DOMAIN	28..193	FIRST LUMENAL DOMAIN.
FT	DOMAIN	194..226	HINGE.
FT	DOMAIN	227..381	SECOND LUMENAL DOMAIN.
FT	DISULFID	40..79	
FT	DISULFID	154..190	
FT	DISULFID	230..268	
FT	DISULFID	337..374	
FT	CARBOHYD	36..36	N-LINKED (GLCNAC. .)
FT	CARBOHYD	44..44	N-LINKED (GLCNAC. .)
FT	CARBOHYD	61..61	N-LINKED (GLCNAC. .)
FT			(POLYLACTOSAMINOGLYCAN).
FT	CARBOHYD	75..75	N-LINKED (GLCNAC. .)
FT	CARBOHYD	83..83	N-LINKED (GLCNAC. .)
FT	CARBOHYD	102..102	N-LINKED (GLCNAC. .)
FT	CARBOHYD	106..106	N-LINKED (GLCNAC. .)
FT	CARBOHYD	120..120	N-LINKED (GLCNAC. .)

```

FT CARBOHYD 129 129 (POLY(LACTOSAMINOGLYCAN).
FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .)
FT CARBOHYD 180 180 (POLY(LACTOSAMINOGLYCAN).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .)
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (PARTIAL).
FT CARBOHYD 199 199 O-LINKED (GALNAC. . .)
FT CARBOHYD 206 206 O-LINKED (GALNAC. . .)
FT CARBOHYD 208 208 O-LINKED (GALNAC. . .)
FT CARBOHYD 210 210 O-LINKED (GALNAC. . .)
FT CARBOHYD 222 222 N-LINKED (GLCNAC. . .)
FT CARBOHYD 227 227 (POLY(LACTOSAMINOGLYCAN).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .)
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .)
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .)
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .)
FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .)
FT CARBOHYD 379 379 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 32 39 VNKGNGTA -> MARGGRVR (IN REF. 3).
SQ SEQUENCE 416 AA; 44773 MW; AAEED793F7A560CB7 CRC64;

```

```

Query Match 76.9%; Score 30; DB 1; Length 416;
Best local similarity 66.7%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 GVNXSSAF 9
DB 290 GVNASSSRF 298

```

```

RESULT 4
UZIP DROME STANDARD; PRT; 488 AA.
ID UZIP DROME
AC P10379; Q960B4; Q9MOW7;
DT 01-MAR-1989 (Rel. 10, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Unzipper protein precursor (zipper protein).
GN UZIP OR ZIP OR CG3533.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Empidoidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88296414; PubMed=3402433;
RA Zhao D.-B., Code S., Jaehning F., Haller J., Jaechle H.;
RT "zipper encodes a putative integral membrane protein required for
RT normal axon patterning during Drosophila neurogenesis.";
RL EMO J. 7.1115-1119(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkeley;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blaise R.G., Champs M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolhakov S.,
RA Borkova D., Botchan M.A., Butler H., Brokstein P., Brotter P.,
RA Burris K.C., Busan D.A., Butler J., Brokstein P., Cantor A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,

```

```

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Urali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matzel B., McIntosh T.C., Kobayashi C., Morris J., Moshireff A.,
RA Merkulov G., Milshina N.V., Moberg C., Mottis J., Moshireff A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusseken D.R., Pacleby J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sienkiewicz I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spadling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng J.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:1285-1295 (2000).
RN [3]
RP REVISIONS.
RX STRAIN=Berkeley; PubMed=12537572;
RX MEDLINE=22426069; PubMed=12537569;
RA Mitra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Belencourt B.R., Celinker S.E., de Grey A.D.N.J., Dysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkeley; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champs M.,
RA George R.A., Garin H., Krommiller B., Pacleby J.M., Park S., Wan K.H.,
RA Rubin G.M., Celinker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -1- FUNCTION: Required for normal axon patterning during neurogenesis.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07450; CAA3032.1; -
DR EMBL; AB003465; AAF47312.2; -
DR EMBL; AY052139; AAK93563.1; -
DR PIR; S00483; EPPF.
DR FLYBase; FBgn0004055; zipper.
KW Differentiation; Neurogenesis; Transmembrane; Signal.
FT SIGNAL 1 21
FT CHAIN 22 488
FT DOMAIN 22 465
FT TRANSMEM 466 486
FT DOMAIN 487 488
FT DOMAIN 380 400
FT CARBOHYD 35 35
FT CARBOHYD 232 232
FT CARBOHYD 317 317
FT CARBOHYD 374 374
FT CARBOHYD 374 374
FT CARBOHYD 374 374

```

```

FT CARBOHYD 448 448 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 488 488 L->SUDEHELMONYRY (IN REF. 1).
SQ SEQUENCE 488 AA, 54308 MW, 58493D9012A2BE7 CRC64;

Query Match
Best Local Similarity 76.9%; Score 30; DB 1; Length 488;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GYNAXSAP 9
DB 463 GYNAAGSTR 471

RESULT 5
CX33 HUMAN STANDARD; PRT; 270 AA.
AC 075712;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Gap junction beta-3 protein (Connexin 31) (CX31).
GN GJB3 OR CX31.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX [1] TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=98369849; PubMed=9704026;
RA Wenzel K., Manthey D., Willecke K., Grzeschik K.-H., Traub O.,
RT "Human gap junction protein connexin31: molecular cloning and
RT expression analysis."
RL Biochem. Biophys. Res. Commun. 248:910-915 (1998).
RN 2
RP SEQUENCE FROM N.A.
RC XIA J.-H., Pan Q., Liu C.-Y., Zheng D., Xie W.;
RT "Molecular cloning of human connexin 31 and 31.1 genes."
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN 3
RP SEQUENCE FROM N.A.
RC Pearce A.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN 141
RP SEQUENCE FROM N.A. AND VARIANTS EXV ASP-12; ARG-12 AND SER-86.
RX MEDLINE=99057343; PubMed=9843209;
RA Richard G., Smith L.E., Bailey R.A., Itin P., Hohl D.,
RA Bostein E.H. Jr., Digiovanna J.J., Compion J.G., Bale S.J.;
RT "Mutations in the human connexin gene GJB3 cause erythrokeratoderma
RT variabilis."
RL Nat. Genet. 20:366-369 (1998).
RN 15
RP SEQUENCE FROM N.A.
RC TISSUE=Urinary bladder;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Deng J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toehlyuk S., Abramson R.D., Mullaly S.J.,
RA Raba S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay J.J., Hulyk S.W.,
RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Morris W.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN 161
RP VARIANT DFN2 LYS-153.
RX MEDLINE=99057344; PubMed=9843210;
RA Xia J.-H., Liu C.-Y., Tang B.S., Pan Q., Huang L., Dai H.P.,
RA Zhang B.R., Xie W., Hu D.X., Zheng D., Shi X.L., Wang D.A., Xia K.,
RA Yu K.P., Liao X.D., Feng Y., Yang Y.F., Xiao J.Y., Xie D.H.,
RA Huang J.Z.;
RT "Mutations in the gene encoding gap junction protein beta-3 associated
RT with autosomal dominant hearing impairment."
RL Nat. Genet. 20:370-373 (1998).
RN 17
RP VARIANT EXV PRO-42.
RX MEDLINE=20062991; PubMed=10594760;
RA Wilgoss A., Leigh I.M., Barnes M.R., Dopping-Hepenstal P.,
RA Eady R.A.U., Walter J.M., Kennedy C.T., Kelsell D.P.;
RT "Identification of a novel mutation R42P in the gap junction protein
RT beta-3 associated with autosomal dominant erythrokeratoderma
RT variabilis."
RL J. Invest. Dermatol. 113:1119-1122 (1999).
RN 18
RP VARIANTS EXV PRO-42 AND LEU-137.
RX MEDLINE=20256220; PubMed=10798362;
RA Richard G., Brown N., Smith L.E., Terrinoni A., Melino G.,
RA Mackie R.M., Bale S.J., Uitto J.;
RT "The spectrum of mutations in erythrokeratodermias -- novel and de
RT novo mutations in GJB3."
RL Hum. Genet. 106:321-329 (2000).
RN 19
RP FUNCTION: One gap junction consists of a cluster of closely packed
RN pairs of transmembrane channels, the connexons, through which
RN materials of low mw diffuse from one cell to a neighboring cell.
RN 20
RP SUBUNIT: A connexon is composed of a hexamer of connexins.
RN 21
RP SUBCELLULAR LOCATION: Integral membrane protein.
RN 22
RP DISEASE: Defects in GJB3 are a cause of erythrokeratoderma
RN variabilis (BKV) [MIM:133300], an autosomal dominant disease
RN characterized by 2 morphologic features: transient figurate red
RN patches and localized or generalized hyperkeratosis.
RN 23
RP DISEASE: Defects in GJB3 are a cause of autosomal dominant non-
RN syndromic sensorineural deafness type 2 (DFNA2) [MIM:600101].
RN 24
RP SIMILARITY: BELONGS TO THE CONNEXIN FAMILY. BETA-TYPE (GROUP I)
RN SUBFAMILY.
RN 25
RP DATABASE: NAME=Connexin-deafness homepage;
RN WWW="http://www.crg.es/deafness/";
RN 26
RP DATABASE: NAME=Hereditary hearing loss homepage;
RN NOTE=Gene page;
RN WWW="http://www.uia.ac.be/dnalab/hhh/".
RN 27
RP THIS SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outbreak -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (see http://www.isb-sib.ch/announce/
RN or send an email to license@isb-sib.ch).
RN 28
RP EMBL: AJ004856; CAA06165.1; -
RN EMBL: AF052692; AAD11815.1; -
RN EMBL: AL121988; CAB90263.1; -
RN EMBL: AF095730; AAC95471.1; -
RN EMBL: BC012918; AAI12918.1; -
RN PIR: J60274; J60274.
RN 29
RP GENE: HGNC:4285; GJB3.
RN 30
RP MIM: 603324; -
RN MIM: 133200; -
RN MIM: 600101; -
RN 31
RP GO: GO:0007154; P:cell communication, NAS.
RN 32
RP InterPro: IPR000500; Connexin.
RN 33
RP Pfam: PF00029; connexin; 1.
RN 34
RP PRINTS: PF00026; CONNEXIN.
RN 35
RP SMART: SM00037; CNX; 1.
RN 36
RP PROSITE: PS00407; CONNEXIN_1; 1.
RN PROSITE: PS00408; CONNEXIN_2; 1.

```



```

Gap junction: Transmembrane; Disease mutation; Deafness; Polymorphism.
FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 21 40 POTENTIAL.
FT DOMAIN 41 75 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 76 98 POTENTIAL.
FT DOMAIN 99 126 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 127 149 POTENTIAL.
FT DOMAIN 150 187 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 188 210 POTENTIAL.
FT DOMAIN 211 270 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 12 12 G->D (in EXV).
FT VARIANT 12 12 G->R (in EXV).
FT TRANSSEM 12 12 G->W (in dbSNF:1805063).
FT VARIANT 32 32 R->P (in EXV).
FT TRANSSEM 42 42 R->S (in EXV).
FT VARIANT 86 86 C->S (in EXV).
FT TRANSSEM 137 137 F->L (in EXV).
FT VARIANT 141 141 FTTD=VAR_015086.
FT TRANSSEM 183 183 FTTD=VAR_015087.
FT VARIANT 183 183 E->K (in DENA2).
FT TRANSSEM 270 AA; 30818 MW; E46D36E5835646A4 CRC64;
SQ SEQUENCE 270 AA; 30818 MW; E46D36E5835646A4 CRC64;

Query Match 74.4%; Score 29; DB 1; Length 270;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVNAXSSAF 9
Db 12 GVNKYSTAF 20

RESULT 6
CXB3 MOUSE STANDARD; PRT; 270 AA.
ID CXB3 MOUSE STANDARD; PRT; 270 AA.
AC P28231;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gap junction beta-3 protein (Connexin 31) (Cx31).
GN GJB3 OR CXN-31.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92347372; PubMed=1322300;
RA Hennemann H., Schwarz H.J., Willecke K.;
RT "Characterization of gap junction genes expressed in P9 embryonic
RT carcinoma cells: molecular cloning of mouse connexin31 and -45
RT cDNAs"; Cell Biol. 57:51-58(1992).
RL Eur. J. Cell Biol. 57:51-58(1992).
CC -1- FUNCTION: One gap junction consists of a cluster of closely packed
CC pairs of transmembrane channels, the connexons, through which
CC materials of low mw diffuse from one cell to a neighboring cell.
CC -1- SUBUNIT: A connexon is composed of a hexamer of connexins.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE CONNEXIN FAMILY. BETA-TYPE (GROUP 1)
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; X63099; CAA44813.1; -.
CC PIR; S23590; S23590.
CC MGD; MGI:95721; Gjb3.
CC InterPro; IPRO00500; Connexin.
CC Pfam; PF00029; connexin; 1.
CC PRINTS; PR00206; CONNEXIN.
CC SMART; SM00037; CNX; 1.
CC PROSITE; PS00407; CONNEXIN_1; 1.
CC PROSITE; PS00408; CONNEXIN_2; 1.
CC Gap junction: Transmembrane.
FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 21 40 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 41 75 POTENTIAL.
FT TRANSSEM 76 98 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 99 126 POTENTIAL.
FT TRANSSEM 127 149 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 150 188 POTENTIAL.
FT TRANSSEM 189 211 POTENTIAL.
FT DOMAIN 212 270 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 270 AA; 30902 MW; A91AB784E03B5869 CRC64;

Query Match 74.4%; Score 29; DB 1; Length 270;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GVNAXSSAF 9
Db 12 GVNKYSTAF 20

RESULT 7
CXB3 RAT STANDARD; PRT; 270 AA.
ID CXB3 RAT STANDARD; PRT; 270 AA.
AC P25305;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gap junction beta-3 protein (Connexin 31) (Cx31).
GN GJB3 OR CXN-31.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91177912; PubMed=1706719;
RA Hoh J.H., John S.A., Revel J.-P.;
RT "Molecular cloning and characterization of a new member of the gap
RT junction gene family, connexin-31";
RL J. Biol. Chem. 266:6524-6531(1991).
CC -1- FUNCTION: One gap junction consists of a cluster of closely packed
CC pairs of transmembrane channels, the connexons, through which
CC materials of low mw diffuse from one cell to a neighboring cell.
CC -1- SUBUNIT: A connexon is composed of a hexamer of connexins.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE CONNEXIN FAMILY. BETA-TYPE (GROUP 1)
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M59336; AAA40997.1; -.
CC PIR; A38737; A38737.
CC InterPro; IPRO00500; Connexin.
CC Pfam; PF00029; connexin; 1.
CC PRINTS; PR00206; CONNEXIN.
CC SMART; SM00037; CNX; 1.

```

```

DR PROSITE; PS00407; CONNEXINS_1; 1.
KM Gap junction; Transmembrane.
FT DOMAIN 1 20
FT TRANSMEM 21 40
FT DOMAIN 41 75
FT TRANSMEM 76 98
FT DOMAIN 99 126
FT TRANSMEM 127 149
FT DOMAIN 150 188
FT TRANSMEM 189 211
FT DOMAIN 212 270
SQ SEQUENCE 270 AA; 30969 MW; 3D0C506FCE179C44 CRC64;

Query Match
Best Local Similarity 74.4%; Score 29; DB 1; Length 270;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYNAXSAP 9
DB 12 GYNKYSTAF 20

RESULT 8
CXBS_MOUSE
ID CXBS_MOUSE STANDARD; PRT; 271 AA.
AC Q02739;
DT 01-FEB-1994 (Rel. 28; Created)
DT 01-FEB-1994 (Rel. 28; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Gap junction beta-5 protein (Connexin 31.1) (Cx31.1).
GN GJB5 OR CXN-31.1
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP MEDLINE=92381038; PubMed=1512260;
RA Hennemann H., Dahl E., White C.B., Schwarz H.J., Lailey P.A.,
RA Chang S., Nicholson B.U., Willecke K.;
RT "Two gap junction genes, connexin 31.1 and 30.3, are closely linked
RT on mouse chromosome 4 and preferentially expressed in skin.";
RL J. Biol. Chem. 267:17225-17233(1992).
CC - FUNCTION: One gap junction consists of a cluster of closely packed
CC pairs of transmembrane channels, the connexons, through which
CC materials of low mw diffuse from one cell to a neighboring cell.
CC - SUBUNIT: A connexon is composed of a hexamer of connexins.
CC - SUBCELLULAR LOCATION: Integral membrane protein.
CC - TISSUE SPECIFICITY: Expressed in skin, much lower in testis.
CC - SIMILARITY: BELONGS TO THE CONNEXIN FAMILY. BETA-TYPE (GROUP I)
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M91236; AAA37426.1; -
DR EMBL; M91442; AAA37427.1; -
DR PIR; B43433; B43433.
DR MGD; MG1:95723; G3b5.
DR InterPro; IPR000500; Connexin.
DR Pfam; PF00029; connexin.1.
DR PRINTS; SM00037; CNX.1.
DR SMART; SM0037; CNX.1.
DR PROSITE; PS00407; CONNEXINS_1; 1.
DR PROSITE; PS00408; CONNEXINS_2; 1.
KM Gap junction; Transmembrane.
FT DOMAIN 1 20
CYTOPLASMIC (POTENTIAL).

```

```

FT TRANSMEM 21 40
FT DOMAIN 41 75
FT TRANSMEM 76 98
FT DOMAIN 99 124
FT TRANSMEM 125 147
FT DOMAIN 148 183
FT TRANSMEM 184 206
FT DOMAIN 207 271
SQ SEQUENCE 271 AA; 31194 MW; F8732F5A8F5CD51A CRC64;

Query Match
Best Local Similarity 74.4%; Score 29; DB 1; Length 271;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYNAXSAP 9
DB 12 GYNKYSTAF 20

RESULT 9
CXBS_RAT
ID CXBS_RAT STANDARD; PRT; 271 AA.
AC P28232;
DT 01-DEC-1992 (Rel. 24; Created)
DT 01-DEC-1992 (Rel. 24; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Gap junction beta-5 protein (Connexin 31.1) (Cx31.1).
GN GJB5 OR CXN-31.1
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP MEDLINE=92112940; PubMed=1370487;
RA Haefliger J.-A., Bruzzone R., Jenkins N.A., Gilbert D.J.,
RA Copeland N.G., Paul D.U.;
RT "Four novel members of the connexin family of gap junction proteins.
RT Molecular cloning, expression, and chromosome mapping.";
RL J. Biol. Chem. 267:2057-2064(1992).
CC - FUNCTION: One gap junction consists of a cluster of closely packed
CC pairs of transmembrane channels, the connexons, through which
CC materials of low mw diffuse from one cell to a neighboring cell.
CC - SUBUNIT: A connexon is composed of a hexamer of connexins.
CC - SUBCELLULAR LOCATION: Integral membrane protein.
CC - TISSUE SPECIFICITY: Expressed in skin.
CC - SIMILARITY: BELONGS TO THE CONNEXIN FAMILY. BETA-TYPE (GROUP I)
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M76533; AAB38538.1; -
DR PIR; D42053; D42053.
DR InterPro; IPR000500; Connexin.
DR Pfam; PF00029; connexin.1.
DR PRINTS; SM00037; CNX.1.
DR SMART; SM0037; CNX.1.
DR PROSITE; PS00407; CONNEXINS_1; 1.
DR PROSITE; PS00408; CONNEXINS_2; 1.
KM Gap junction; Transmembrane.
FT DOMAIN 1 20
FT TRANSMEM 21 40
FT DOMAIN 41 75
FT TRANSMEM 76 98
FT DOMAIN 99 124
FT TRANSMEM 125 147
FT DOMAIN 148 182
CYTOPLASMIC (POTENTIAL).
EXTRACELLULAR (POTENTIAL).

```

```

FT TRANSMEM 183 205 POTENTIAL.
SQ DOMAIN 206 271 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 271 AA; 31046 MW; 7753FC96909015CE CRC64;

Query Match
Best Local Similarity 66.7%; Score 29; DB 1; Length 271;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
DB 12 GVNKYSTAF 20

RESULT 10
CXBS_HUMAN STANDARD; PRT; 273 AA.
ID 095377; Q9UPA3;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Gap junction beta-5 protein (Connexin 31.1) (Cx31.1).
GN Cx31.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP MEDLINE=99057343; PubMed=9843209;
RA Richard G., Smith L.E., Bailey R.A., Itin P., Hohl D.,
RA Epstein E.H. Jr., DiGiovanna J.J., Compton J.G., Bale S.J.;
RT "Mutations in the human connexin gene GJB3 cause erythrokeratoderma
RT variabilis";
RL Nat. Genet. 20:366-369(1998).
RN 12
RP SEQUENCE FROM N.A.
RA Xia J.-H., Liu C.-Y., Zheng D., Pan Q., Xie W.;
RT "Molecular cloning of human connexin 31 and 31.1."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN 13
RP SEQUENCE FROM N.A.
RA Pearce A.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN 14
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rata S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Harte S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodigues S., Sanchez A.,
RA Whiting M., Madan A., Kettman M., Madan A., Rodigues S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalins D.E.,
RA Schenker A., Schein J.E., Jones S.U.M., Maitz M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: One gap junction consists of a cluster of closely packed
CC pairs of transmembrane channels, the connexons, through which
CC materials of low mw diffuse from one cell to a neighboring cell.
CC -1- SUBUNIT: A connexon is composed of a hexamer of connexins.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE CONNEXIN FAMILY. BETA-TYPE (GROUP I)

```

```

CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF099731; AAC95472.1; -
DR EMBL; AF052693; AAD18005.1; -
DR EMBL; AL121988; CAB90271.1; -
DR EMBL; BC004379; AAH04379.1; -
DR Genem; HGNC:4267; GJB5.
DR MIM; 604493; -
DR GO; GO:0008544; P:epidermal differentiation; TAS.
DR InterPro; IPR000500; Connexin.
DR Pfam; PF00029; connexin; 1.
DR PRINTS; PR00206; CONNEXIN.
DR SMART; SM00037; CNX; 1.
DR PROSITE; PS00407; CONNEXIN_1; 1.
DR PROSITE; PS00408; CONNEXIN_2; 1.
KW Gap junction; Transmembrane.
FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 21 40 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 41 75 POTENTIAL.
FT TRANSMEM 76 98 POTENTIAL.
FT TRANSMEM 99 126 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 127 149 POTENTIAL.
FT DOMAIN 150 187 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 188 210 POTENTIAL.
FT DOMAIN 211 273 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 273 AA; 31088 MW; C1B9F1AA42E070B8 CRC64;

Query Match
Best Local Similarity 74.4%; Score 29; DB 1; Length 273;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
DB 12 GVNKYSTAF 20

RESULT 11
LMAS_MOUSE STANDARD; PRT; 3718 AA.
ID 061051; Q9JH06;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Laminin alpha-5 chain precursor.
GN LAMA5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE OF 1-92 FROM N.A. AND SEQUENCE OF 41-46.
RX MEDLINE=21818471; PubMed=11829758;
RA Garbe J.H., Gohring W., Mann K., Timpl R., Sasaki T.;
RT "Complete sequence, recombinant analysis and binding to laminins and
RT alpha5 chains of the N-terminal domains of laminin alpha3b and
RT alpha5 chains";
RL Biochem. J. 362:213-221(2002).
RN 12
RP SEQUENCE OF 84-3718 FROM N.A.
RX STRAIN=C57BL/6 X CBA; TISSUE=Lung;
RX MEDLINE=96081906; PubMed=7499364;
RA Miner J.H., Lewis R.M., Sames J.R.;
RT "Molecular cloning of a novel laminin chain, alpha 5, and widespread
RT expression in adult mouse tissues";
RL J. Biol. Chem. 270:28523-28526(1995).

```

RN [3]
 RP REVISIONS.
 RA Miner J.H., Lewis R.M., Sanes J.R.;
 RL Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
 is thought to mediate the attachment, migration, and organization
 of cells into tissues during embryonic development by interacting
 with other extracellular matrix components.
 CC -1- FUNCTION: ALPHA-5 CHAIN MAY BE THE MAJOR LAMININ ALPHA CHAIN OF
 ADULT EPITHELIAL AND/OR ENDOTHELIAL BASAL LAMINAE.
 CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
 different polypeptide chains (alpha, beta, gamma), which are bound
 to each other by disulfide bonds into a cross-shaped molecule
 comprising one long and three short arms with globules at each
 end.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
 MEMBRANES (MAJOR COMPONENT).
 CC -1- TISSUE SPECIFICITY: IN ADULT, HIGH LEVELS IN HEART, LUNG, AND
 KIDNEY; LOWER IN BRAIN, MUSCLE AND TESTIS; VERY LOW IN LIVER, GUT
 AND SKIN. EXPRESSED IN MANY TISSUES IN EMBRYONIC DAY 11.
 CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC -1- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
 CC -1- SIMILARITY: Contains 1 laminin N-terminal domain.
 CC -1- SIMILARITY: Contains 22 laminin EGF-like domains.
 CC -1- SIMILARITY: Contains 5 laminin G-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ293593; CAB9255.1; -
 DR EMBL; U37501; AAC53430.1; -
 DR PIR; T10053; T10053.
 DR HSSP; P02468; 1TLE.
 DR MGD; MG1105382; Lama5.
 DR GO; GO:0005604; C:Basement membrane; IDA.
 DR GO; GO:0030324; P:Lung development; IMP.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR000034; Laminin B.
 DR InterPro; IPR002049; Laminin_EGF.
 DR InterPro; IPR001791; Laminin_G.
 DR InterPro; IPR001886; LAMNT.
 DR Pfam; PF00052; Laminin_B; 1.
 DR Pfam; PF00053; Laminin_EGF; 18.
 DR Pfam; PF00054; Laminin_G; 2.
 DR Pfam; PF00055; Laminin_Nterm; 1.
 DR PRINTS; PR00011; EGF-LAMININ.
 DR ProDom; PD002082; Lam_N2; 1.
 DR ProDom; PD003031; Laminin_B; 1.
 DR SMART; SK00180; EGF_Lam; 20.
 DR SMART; SK00281; LamB; 1.
 DR SMART; SK00282; LamG; 5.
 DR SMART; SM00136; LAMNT; 1.
 DR PROSITE; PS00022; EGF_1; 19.
 DR PROSITE; PS01186; EGF_2; 3.
 DR PROSITE; PS01248; LAMININ_TYPE_EGF; 19.
 DR PROSITE; PS50028; LAM_G_DOMAIN; 5.
 KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
 FT SIGNAL
 FT CHAIN 1 40
 FT 1 3718 LAMININ ALPHA-5 CHAIN.
 FT DOMAIN 41 304 LAMININ N-TERMINAL (DOMAIN VI).
 FT 305 363 LAMININ EGF-LIKE 1.
 FT DOMAIN 364 433 LAMININ EGF-LIKE 2.
 FT DOMAIN 434 479 LAMININ EGF-LIKE 3.
 FT DOMAIN 500 546 LAMININ EGF-LIKE 4.
 FT DOMAIN 547 592 LAMININ EGF-LIKE 5.

FT DOMAIN 593 637 LAMININ EGF-LIKE 6.
 FT DOMAIN 638 682 LAMININ EGF-LIKE 7.
 FT DOMAIN 683 728 LAMININ EGF-LIKE 8.
 FT DOMAIN 729 781 LAMININ EGF-LIKE 9.
 FT DOMAIN 782 833 LAMININ EGF-LIKE 10.
 FT DOMAIN 834 855 LAMININ EGF-LIKE 11 (INCOMPLETE).
 FT DOMAIN 856 1442 LAMININ DOMAIN IV 1 (DOMAIN IV B).
 FT DOMAIN 1443 1488 LAMININ EGF-LIKE 12.
 FT DOMAIN 1489 1532 LAMININ EGF-LIKE 13.
 FT DOMAIN 1533 1581 LAMININ EGF-LIKE 14.
 FT DOMAIN 1582 1632 LAMININ EGF-LIKE 15.
 FT DOMAIN 1633 1642 LAMININ EGF-LIKE 16 (N-TERMINAL).
 FT DOMAIN 1643 1831 LAMININ DOMAIN IV 2 (DOMAIN IV A).
 FT DOMAIN 1832 1864 LAMININ EGF-LIKE 16 (C-TERMINAL).
 FT DOMAIN 1864 1914 LAMININ EGF-LIKE 17.
 FT DOMAIN 1915 1958 LAMININ EGF-LIKE 18.
 FT DOMAIN 1959 1970 LAMININ EGF-LIKE 18.
 FT DOMAIN 1971 2024 LAMININ EGF-LIKE 19.
 FT DOMAIN 2025 2071 LAMININ EGF-LIKE 20.
 FT DOMAIN 2072 2118 LAMININ EGF-LIKE 21.
 FT DOMAIN 2119 2168 LAMININ EGF-LIKE 22.
 FT DOMAIN 2169 2735 DOMAIN II AND I.
 FT DOMAIN 2736 2933 LAMININ G-LIKE 1.
 FT DOMAIN 2934 3119 LAMININ G-LIKE 2.
 FT DOMAIN 3128 3296 LAMININ G-LIKE 3.
 FT DOMAIN 3337 3511 LAMININ G-LIKE 4.
 FT DOMAIN 3518 3689 LAMININ G-LIKE 5.
 FT DOMAIN 2205 2257 COILED COIL (POTENTIAL).
 FT DOMAIN 2330 2461 COILED COIL (POTENTIAL).
 FT DOMAIN 2604 2621 COILED COIL (POTENTIAL).
 FT DOMAIN 2639 2705 COILED COIL (POTENTIAL).
 FT SITE 1733 1725 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 1839 1841 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 305 314 BY SIMILARITY.
 FT DISULFID 307 327 BY SIMILARITY.
 FT DISULFID 329 338 BY SIMILARITY.
 FT DISULFID 341 361 BY SIMILARITY.
 FT DISULFID 364 373 BY SIMILARITY.
 FT DISULFID 366 398 BY SIMILARITY.
 FT DISULFID 401 410 BY SIMILARITY.
 FT DISULFID 413 431 BY SIMILARITY.
 FT DISULFID 434 445 BY SIMILARITY.
 FT DISULFID 436 452 BY SIMILARITY.
 FT DISULFID 454 463 BY SIMILARITY.
 FT DISULFID 466 476 BY SIMILARITY.
 FT DISULFID 500 512 BY SIMILARITY.
 FT DISULFID 502 521 BY SIMILARITY.
 FT DISULFID 523 532 BY SIMILARITY.
 FT DISULFID 535 544 BY SIMILARITY.
 FT DISULFID 547 559 BY SIMILARITY.
 FT DISULFID 568 577 BY SIMILARITY.
 FT DISULFID 568 590 BY SIMILARITY.
 FT DISULFID 593 605 BY SIMILARITY.
 FT DISULFID 595 611 BY SIMILARITY.
 FT DISULFID 613 622 BY SIMILARITY.
 FT DISULFID 625 635 BY SIMILARITY.
 FT DISULFID 638 650 BY SIMILARITY.
 FT DISULFID 640 656 BY SIMILARITY.
 FT DISULFID 658 667 BY SIMILARITY.
 FT DISULFID 670 680 BY SIMILARITY.
 FT DISULFID 683 695 BY SIMILARITY.
 FT DISULFID 685 702 BY SIMILARITY.
 FT DISULFID 704 713 BY SIMILARITY.
 FT DISULFID 716 726 BY SIMILARITY.
 FT DISULFID 1443 1455 BY SIMILARITY.
 FT DISULFID 1445 1462 BY SIMILARITY.
 FT DISULFID 1464 1473 BY SIMILARITY.
 FT DISULFID 1476 1486 BY SIMILARITY.
 FT DISULFID 1533 1548 BY SIMILARITY.
 FT DISULFID 1535 1555 BY SIMILARITY.
 FT DISULFID 1557 1566 BY SIMILARITY.
 FT DISULFID 1569 1579 BY SIMILARITY.
 FT DISULFID 1582 1594 BY SIMILARITY.

```

FT DISULFID 1584 1601 BY SIMILARITY.
FT DISULFID 1603 1612 BY SIMILARITY.
FT DISULFID 1615 1630 BY SIMILARITY.
FT DISULFID 1865 1874 BY SIMILARITY.
FT DISULFID 1867 1891 BY SIMILARITY.
FT DISULFID 1884 1893 BY SIMILARITY.
FT DISULFID 1896 1912 BY SIMILARITY.
FT DISULFID 1915 1930 BY SIMILARITY.
FT DISULFID 1917 1939 BY SIMILARITY.
FT DISULFID 1941 1950 BY SIMILARITY.
FT DISULFID 1953 1968 BY SIMILARITY.
FT DISULFID 1971 1985 BY SIMILARITY.
FT DISULFID 1973 1993 BY SIMILARITY.
FT DISULFID 1996 2005 BY SIMILARITY.
FT DISULFID 2008 2022 BY SIMILARITY.
FT DISULFID 2072 2083 BY SIMILARITY.
FT DISULFID 2074 2090 BY SIMILARITY.
FT DISULFID 2092 2101 BY SIMILARITY.
FT DISULFID 2104 2116 BY SIMILARITY.
FT DISULFID 2119 2126 BY SIMILARITY.
FT DISULFID 2121 2133 BY SIMILARITY.
FT DISULFID 2135 2144 BY SIMILARITY.
FT DISULFID 2147 2166 BY SIMILARITY.
FT DISULFID 2169 2172 INTERCHAIN (PROBABLE).
FT DISULFID 2172 INTERCHAIN (PROBABLE).
FT CARBOHYD 100 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 148 N-LINKED (GLCNAC...) (POTENTIAL).

```

Query Match 74.4%; Score 29; DB 1; Length 3718;
 Best Local Similarity 55.6%; Pred. No. 3e-02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYNXSSAF 9
 Db 2196 GYNXSSAF 2204

```

RESULT 12
YEVL YEAST STANDARD; PRT; 114 AA.
AC P40076;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 12.9 kDa protein in SC92-G103 intergenic region.
GN YER121W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RX PubMed=9169868;
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Yelton M.A., Allen E.,
RA Arango R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA Huntick-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA Lin D., Mesedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA Smith V., Taylor P., Wei Y., Borstein D., Davis R.W.;
RL "The nucleotide sequence of Saccharomyces cerevisiae chromosome V."
RL Nature 387:78-81(1997).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; U18916; AAC03219.1; -

```

```

DR PIR; S50624; S50624.
DR SGD; S0000923; YER121W.
KW Hypothetical protein.
SQ SEQUENCE 114 AA; 12941 MW; 71AEDD8607BE774B CRC64;

```

Query Match 71.8%; Score 28; DB 1; Length 114;
 Best Local Similarity 85.7%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 NAXSSAF 9
 Db 90 NAXSSAF 96

```

RESULT 13
TATD_ECOLI STANDARD; PRT; 260 AA.
AC P27859; P27860; P78128; P78129;
DT 01-AUG-1992 (Rel. 23, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Deoxyribonuclease (acid) (EC 3.1.21.-) (DNase (acid)).
GN TATD OR MTC OR B3840/B3841.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=92358234; PubMed=1379743;
RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
RT "Analysis of the Escherichia coli genome: DNA sequence of the region
RT from 84.5 to 86.5 minutes."
RL Science 257:771-778 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=5Kc;
RX MEDLINE=92261317; PubMed=1584020;
RA Bailey M.J.A., Koronakis V., Schmolli T., Hughes C.;
RT "Escherichia coli HlyT protein, a transcriptional activator of
RT hemolysin synthesis and secretion, is encoded by the hlyA (stxB)
RT locus required for expression of sex factor and lipopolysaccharide
RT genes."
RL Mol. Microbiol. 6:1003-1021(1992).
RN [3]
RP SEQUENCE FROM N.A., AND GENE NAME.
RC STRAIN=K12;
RX MEDLINE=98315056; PubMed=9649434;
RA Sargent F., Bogsch E.G., Stanley N.R., Wexler M., Robinson C.,
RA Berks B.C., Palmer T.;
RT "Overlapping functions of components of a bacterial Sec-independent
RT protein export pathway."
RL EMBO J. 17:3640-3650(1998).
RN [4]
RP PARTIAL SEQUENCE OF 1-6, AND CHARACTERIZATION.
RX MEDLINE=20287539; PubMed=10747959;
RA Wexler M., Sargent F., Jack R.L., Stanley N.R., Bogsch E.G.,
RA Robinson C., Berks B.C., Palmer T.;
RT "Tad is a cytoplasmic protein with DNase activity. No requirement for
RT Tad family proteins in sec-independent protein export."
RL J. Biol. Chem. 275:16717-16722(2000).
CC -1- FUNCTION: SHOWS DNASE ACTIVITY.
CC -1- COPACTOR: Magnesium.
CC -1- SUBUNIT: Monomer (Probable).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE TAD DNASE FAMILY.
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT ERROR.
CC -1- CAUTION: WAS ORIGINALLY (REF.4) THOUGHT TO PARTICIPATE IN THE SEC-
CC INDEPENDENT PROTEIN TRANSLOCASE SYSTEM.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M87049; AAA67636.1; ALT_FRAME.
 DR EMBL; M87049; AAA67637.1; ALT_FRAME.
 DR EMBL; AE000459; AAC76843.1; ALT_FRAME.
 DR EMBL; AE000459; AAC76844.1; ALT_FRAME.
 DR EMBL; X65013; -; NOT_ANNOTATED_CDS.
 DR EMBL; AU005830; CAA06727.1; ALT_INIT.
 DR EcoGene; Egl1481; taclD.
 DR InterPro; IPR001130; TaclD_Nasef_1.
 DR Pfam; PF01026; TaclD_Nasef_1.
 DR PROSITE; PS01137; TARD_1; FALSE_NEG.
 DR PROSITE; PS01080; TARD_2; 1.
 DR PROSITE; PS01091; TARD_3; 1.
 KW Hydrolase; Nuclease; Magnesium; Complete proteome.
 SQ SEQUENCE 260 AA; 28974 MW; C18FB5269849746F CRC64;

Query Match 71.8%; Score 28; DB 1; Length 260;
 Best Local Similarity 66.7%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
 Db 5 GVNLTSSQF 13

RESULT 14
 YK03_CAEEL STANDARD; PRT; 401 AA.
 AC P34291;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Hypothetical 43.5 kDa protein C05B5.3 in chromosome III.
 GN C05B5.3.
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 CC Rhabditidae; Peloderinae; Caenorhabditis.
 CC NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Mortimore B.J.;
 RL Submitted (APR-1994) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP REVISIONS.
 RA Durbin R.;
 RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z32679; CAA83596.1; -.
 DR PIR; C88571; C88571.
 DR WormPep; C05B5.3; CE17369.
 KW Hypothetical protein.
 SQ SEQUENCE 401 AA; 43472 MW; C99DF259EF6C8B55 CRC64;

Query Match 71.8%; Score 28; DB 1; Length 401;
 Best Local Similarity 66.7%; Pred. No. 48;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9

Db 239 GVNPTSGAF 247

RESULT 15
 MOCR_RHIME STANDARD; PRT; 493 AA.
 AC P49369;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Probable rhizopine catabolism regulatory protein mocr.
 GN MOCR.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 CC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 CC NCB1_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L5-30;
 RX MEDLINE=95147842; PubMed=784533;
 RA Rosbach S., Kulpa D.A., Rosbach U., de Bruijn F.J.;
 RT "Molecular and genetic characterization of the rhizopine catabolism
 RT (mocr) genes of Rhizobium meliloti U5-30.";
 RL Mol. Gen. Genet. 245:11-24(1994).
 CC -1- FUNCTION: COULD PLAY A REGULATORY ROLE IN THE TRANSCRIPTION OF THE
 CC MOC GENES FOR RHIZOPINE CATABOLISM. COULD ALSO HAVE AN
 CC AMINOTRANSFERASE ACTIVITY.
 CC -1- COPACITOR: Pyridoxal phosphate (potential)
 CC -1- SIMILARITY: BELONGS TO THE GNTT FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC -1- SIMILARITY: IN THE C-TERMINAL, TO AMINOTRANSFERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X78503; CAA55271.1; -.
 DR PIR; S51574; S51574.
 DR InterPro; IPR000524; HTH_GntR.
 DR Pfam; PF00392; GntR_1.
 DR PRINTS; PR00035; HTHGNTT.
 DR SMART; SM00345; HTH_GntR; 1.
 DR PROSITE; PS00043; HTH_GntR_FAMILY; 1.
 KW Transcription regulation; DNA-binding; Transferase; Aminotransferase;
 KW Pyridoxal phosphate.
 FT DNA_BIND 38
 FT BINDING 319
 SQ SEQUENCE 493 AA; 55128 MW; CE1B0FC9C7175C CRC64;

Query Match 71.8%; Score 28; DB 1; Length 493;
 Best Local Similarity 55.6%; Pred. No. 60;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
 Db 464 GINSURSAF 472

Search completed: January 29, 2004, 14:53:11
 Job time: 5.82353 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:46:19 ; Search time 26.8676 Seconds
(without alignments)
86.441 Million cell updates/sec

Title: US-10-032-950-7
Perfect score: 39
Sequence: 1 GYNAXSAF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP yvirus:*
16: SP bacteriaph:*
17: SP archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	82.1	236	2 Q9X494	Q9X494 mycoplasma
2	31	79.5	47	1 Q3586	Q3586 staphylococ
3	31	79.5	924	16 Q97P38	Q97P38 streptococ
4	31	79.5	924	16 Q8CYG6	Q8CYG6 streptococ
5	30	76.9	163	16 Q8CC58	Q8CC58 mycobacteri
6	30	76.9	167	16 Q53195	Q53195 mycobacteri
7	30	76.9	169	2 Q9X7B5	Q9X7B5 mycobacteri
8	30	76.9	189	11 Q922T9	Q922T9 mus musculu
9	30	76.9	214	8 Q8M6L3	Q8M6L3 taenia soli
10	30	76.9	251	4 Q96140	Q96140 homo sapien
11	30	76.9	311	4 Q9BBD2	Q9BBD2 homo sapien
12	30	76.9	355	4 Q9NP13	Q9NP13 mus sapien
13	30	76.9	406	11 Q8YH34	Q8YH34 mus musculu
14	30	76.9	407	11 Q9DC13	Q9DC13 mus musculu
15	30	76.9	417	4 Q8WU33	Q8WU33 homo sapien
16	30	76.9	488	5 Q9WOW7	Q9WOW7 drosophila

17	30	76.9	529	5 Q09593	Q09593 caenorhabdi
18	30	76.9	655	5 Q81RV2	Q81RV2 drosophila
19	30	76.9	659	5 Q8MT98	Q8MT98 drosophila
20	30	76.9	809	5 Q76869	Q76869 drosophila
21	30	76.9	859	5 Q81RV3	Q81RV3 drosophila
22	30	76.9	887	5 Q9W4V1	Q9W4V1 drosophila
23	29	74.4	115	11 P97705	P97705 rattus norv
24	29	74.4	198	16 Q8DSR6	Q8DSR6 streptococ
25	29	74.4	225	2 Q2RFR6	Q2RFR6 rhodobacter
26	29	74.4	243	9 Q8SBQ5	Q8SBQ5 bacterioph
27	29	74.4	266	13 Q8OGJ5	Q8OGJ5 figu rubrip
28	29	74.4	268	13 Q8CFX8	Q8CFX8 xenopus lae
29	29	74.4	271	11 Q91XA4	Q91XA4 mus musculu
30	29	74.4	274	16 Q8CV82	Q8CV82 oceanobacti
31	29	74.4	289	2 Q2XBR8	Q2XBR8 bacillus ce
32	29	74.4	299	13 Q8AWR5	Q8AWR5 brachydanio
33	29	74.4	316	17 Q29728	Q29728 arabidopsi
34	29	74.4	329	16 Q8EYX5	Q8EYX5 leptospira
35	29	74.4	333	16 Q34933	Q34933 bacillus su
36	29	74.4	353	5 Q9GNB2	Q9GNB2 dirofilaria
37	29	74.4	373	2 Q8VTV2	Q8VTV2 staphylococ
38	29	74.4	376	10 Q8S297	Q8S297 oryza sativ
39	29	74.4	411	16 P73214	P73214 synecocyst
40	29	74.4	431	16 Q8U924	Q8U924 agrobacteri
41	29	74.4	463	16 Q8XHH9	Q8XHH9 clostridium
42	29	74.4	506	10 Q8GZS8	Q8GZS8 allium sati
43	29	74.4	591	2 Q9AHY9	Q9AHY9 photorabdu
44	29	74.4	611	10 Q9LRT5	Q9LRT5 arabidopsi
45	29	74.4	611	10 Q8GZT8	Q8GZT8 arabidopsi

ALIGNMENTS

RESULT 1	Q9X494	PRELIMINARY;	PRT;	236 AA.
AC	Q9X494:			
DT	01-NOV-1999 (TREMBLrel. 12, Created)			
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Yabc protein.			
GN	YABC.			
OS	Mycoplasma fermentans.			
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.			
OX	NCBI_TaxId=2115;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=EG18;			
RX	MEDLINE=99115554; PubMed=9916088;			
RA	Calcutt M.V., Kim M.F., Kaipras A.B., Muhlradt P.F., Wise K.S.;			
RT	"Differential posttranslational processing confers intraspecies			
RT	variation of a major surface lipoprotein and a macrophage-activating			
RT	lipopeptide of Mycoplasma fermentans.";			
RL	Infect. Immun. 67:760-771(1999).			
DR	EMBL; AF100324; AAD25734.1.-			
DR	InterPro; IPR000878; CoI/por_Mettransf.			
DR	InterPro; IPR000578; UPF0011.			
DR	Pfam; PF00590; TP_methylase; 1.			
DR	TIGRPFAM; TIGR00096; TIGR00096; 1.			
DR	PROSITE; PS01296; UPF0011; 1.			
SO	SEQUENCE 236 AA; 26620 MW; F129742326E0A35 CRC64;			

Query Match	82.1%;	Score 32;	DB 2;	Length 236;
Best Local Similarity	77.8%;	Pred. No. 29;		
Matches	7;	Conservative	0;	Mismatches
			2;	Indels
				Gaps
				0;
Qy	1 GYNAXSAF 9			
Db	112 GYNAXSAF 120			

RESULT 2

```

O33586
ID O33586 PRELIMINARY; PRT; 47 AA.
AC O33586;
RT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE AGND (AGND protein)
OS AGND OR SAV2037 OR SA1842.1 OR SAS066.
OS Staphylococcus aureus (strain MU50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OC Staphylococcus aureus.
NCBI_TaxID=156878, 156879, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus; STRAIN=SA502A;
RX MEDLINE=97342847; PubMed=9197262;
RA J.G., Beavis R., Novick R.P.;
RT "Bacterial interference caused by autoinducing peptide variants.";
RL Science 276:2027-2030(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain MU50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Karamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyata A.,
RA Mitani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuuchi J.,
RA Kanehisa M., Yamashita A., Oshina K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiratsuka K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AF001782; AA63285.1;
DR EMBL; AP003364; BAB58199.1;
DR EMBL; AP003135; BAB43124.1;
KW Complete proteome.
SQ SEQUENCE 47 AA; 5149 MW; 22893764DF54BFA7 CRC64;

Query Match 79.5%; Score 31; DB 16; Length 47;
Best Local Similarity 77.8%; Pred. No. 8.7;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
DB 24 GVNACSSLF 32

RESULT 3
O97PG8 PRELIMINARY; PRT; 924 AA.
ID O97PG8
AC O97PG8;
RT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
GN Hypothetical protein SP1652.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Petersen H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Petersen S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Knout H., Wolf A.M., Usterack T.R., Hansen C.L.,
RA McDonald L.A., Feldlym T.V., Angiolli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.;

```

```

RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
DR EMBL; AEC07458; AAK75731.1;
DR TIGR; SP1652;
DR InterPro; IPR003838; DUF214.
DR Pfam; PF02687; DUF214; 2.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 924 AA; 102997 MW; C75B27650ED9C65F CRC64;

Query Match 79.5%; Score 31; DB 16; Length 924;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
DB 239 GVNAFSSAY 247

RESULT 4
O8CYG6 PRELIMINARY; PRT; 924 AA.
ID O8CYG6
AC O8CYG6;
RT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Hypothetical protein.
GN SP1496.
OS Streptococcus pneumoniae (strain ATCC BAA-295 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgess S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA Leblanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McShane S.M., McHenry M., Mcleaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
DR EMBL; AF008518; AAI00300.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 924 AA; 102967 MW; 192383CF91548592 CRC64;

Query Match 79.5%; Score 31; DB 16; Length 924;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
DB 239 GVNAFSSAY 247

RESULT 5
O9CC58 PRELIMINARY; PRT; 163 AA.
ID O9CC58
AC O9CC58;
RT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE Hypothetical protein ML1255.
GN ML1255.
OS Mycobacterium lepre.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1769;

```



```

RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=TW;
RX      MEDLINE=2128732; PubMed=1234002;
RA      Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
RA      Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA      Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA      Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA      Holroyd S., Hornsby T., Jagsels K., Lacroix C., Maclean J., Moule S.,
RA      Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA      Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA      Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA      Barrell B.G.;
RT      "Massive gene decay in the leprosy bacillus.";
RL      Nature 408:1007-1011(2001).
DR      EMBL; AL583921; CAC31636.1; -.
DR      Lepioma; ML1255; -.
KW      Hypochemical protein; Complete proteome.
SQ      SEQUENCE 163 AA; 16842 MW; F26A1A2C04E3FD2E CRC64;

Query Match          76.9%; Score 30; DB 16; Length 163;
Best Local Similarity 77.8%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GVNAKSSAF 9
Db      132 GVNAKSSAF 140

RESULT 6
053195 PRELIMINARY; PRT; 167 AA.
AC      053195;
DT      01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT      01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE      01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE      Hypochemical protein RV2468C.
GN      RV2468C OR MTV008.24C OR MT2543.
OS      Mycobacterium tuberculosis.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Corynebacteriinae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1773;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=H37RV;
RX      MEDLINE=98295987; PubMed=9634230;
RA      Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA      Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA      Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA      Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA      Hornsby T., Jagsels K., Krogsh A., McLean J., Moule S., Murphy L.,
RA      Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers R.,
RA      Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA      Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT      "Deciphering the biology of Mycobacterium tuberculosis from the
RT      complete genome sequence.";
RL      Nature 393:537-544(1998).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CDC 1551 / Oshkosh;
RX      MEDLINE=98295987; PubMed=9634230;
RA      Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA      Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickley E.,
RA      Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA      Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA      Bishai W.;
RT      "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT      laboratory strains.";
RL      Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AL021246; CAA16045.1; -.
DR      EMBL; AE007091; AAK46843.1; -.
DR      TIGR; MT2543; -.
DR      TubercuList; RV2468C; -.
KW      Hypochemical protein; Complete proteome.

```

```

SQ      SEQUENCE 167 AA; 17288 MW; F60B6FCA54E4BB4 CRC64;
Query Match          76.9%; Score 30; DB 16; Length 167;
Best Local Similarity 77.8%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GVNAKSSAF 9
Db      136 GVNAKSSAF 144

RESULT 7
09X7B5 PRELIMINARY; PRT; 169 AA.
AC      09X7B5;
DT      01-NOV-1999 (TrEMBLrel. 12, Created)
DT      01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT      01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE      Hypochemical 17.4 kDa protein.
GN      MCB1610.16.
OS      Mycobacterium leprae.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Corynebacteriinae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1769;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Seeger K.J., Harris D.;
RL      Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RA      James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL      Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=93188700; PubMed=8446027;
RA      Eigmeier K., Honore N., Woods S.A., Caudron B., Cole S.T.;
RT      "Use of an ordered cosmid library to deduce the genomic organization
RT      of Mycobacterium leprae.";
RL      Mol. Microbiol. 7:197-206(1993).
DR      EMBL; AL049913; CAB3162.1; -.
KW      Hypochemical protein.
SQ      SEQUENCE 168 AA; 17360 MW; 2FB50C9B845D0B89 CRC64;

Query Match          76.9%; Score 30; DB 2; Length 169;
Best Local Similarity 77.8%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GVNAKSSAF 9
Db      136 GVNAKSSAF 146

RESULT 8
0922T9 PRELIMINARY; PRT; 169 AA.
AC      0922T9;
DT      01-DEC-2001 (TrEMBLrel. 19, Created)
DT      01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Similar to lysosomal membrane glycoprotein 1 (Fragment).
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Strausberg R.;
RL      Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; BC006785; AAH06785.1; -.
DR      InterPro; IPR002000; Lamp.
DR      Pfam; PF01299; Lamp. 1.
DR      PRINTS; PR00336; LYSASSOCTDMP.
DR      PROSITE; PS00310; Lamp_1; 1.

```

DR PROSITE: PS00311; LAMP_2; 1.
 FT NON TER 1
 SQ SEQUENCE 189 AA; 20456 MW; 71F16D65BA4066FA CRC64;

Query Match
 Best Local Similarity 66.7%; Score 30; DB 11; Length 189;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
 Db 63 GVNAXSSAF 71

RESULT 9
 Q9M6L3 PRELIMINARY; PRT; 214 AA.

AC Q9M6L3
 DT 01-OCT-2002 (TRENBLREL. 22, Created)
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE Cytochrome c oxidase subunit 3 (EC 1.9.3.1) (Cytochrome c oxidase polypeptide III).
 GN COX3
 OS Taenia solium.
 OC Microchondrion.
 OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
 OC Cyclophyllidae; Taeniidae; Taenia.
 CX NCBI_TaxID=6204;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nakao M.;
 RT "The complete mitochondrial DNA sequence of the tapeworm Taenia solium."
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: SUBUNITS I, II AND III FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME C + 2 H(2)O.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
 CC EMBL; AB086256; BAC00834.1; -
 DR InterPro; IPR000298; Cytochrome c oxidase III.
 DR ProDom; PD000382; Cytochrome c oxidase III.
 DR PROSITE; PS50253; COX3.1.
 KW Oxidoreductase; Transmembrane; Mitochondrion.
 SQ SEQUENCE 214 AA; 24442 MW; 4E59E1E14E876C5E CRC64;

Query Match
 Best Local Similarity 76.9%; Score 30; DB 8; Length 214;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
 Db 144 GVNAXSSAF 152

RESULT 10
 Q96I40 PRELIMINARY; PRT; 251 AA.

AC Q96I40
 DT 01-DEC-2001 (TRENBLREL. 19, Created)
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
 DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
 DE Hypothetical protein (Fragment).
 GN LAMP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Brain;
 RA Strausberg R.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC007845; AA07845.1; -

DR InterPro; IPR002000; Lamp.
 DR Pfam; PF01289; Lamp_1.
 DR PRINTS; PR00336; LYSASOCTDMP.
 DR PROSITE; PS00310; LAMP_1; 1.
 DR PROSITE; PS00311; LAMP_2; 1.
 KW Hypothetical protein.

FT NON TER 1
 SQ SEQUENCE 251 AA; 26855 MW; 4DA5F87445A3CEB2 CRC64;

Query Match
 Best Local Similarity 76.9%; Score 30; DB 4; Length 251;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
 Db 125 GVNAXSSAF 133

RESULT 11

Q9BRD2 PRELIMINARY; PRT; 311 AA.

AC Q9BRD2
 DT 01-JUN-2001 (TRENBLREL. 17, Created)
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)
 DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
 DE Similar to lysosomal-associated membrane protein 1 (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Brain;
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC006345; AA06345.1; -
 DR InterPro; IPR002000; Lamp.
 DR Pfam; PF01289; Lamp; 1.
 DR PRINTS; PR00336; LYSASOCTDMP.
 DR PROSITE; PS00310; LAMP_1; 1.
 DR PROSITE; PS00311; LAMP_2; 1.
 FT NON TER 1
 SQ SEQUENCE 311 AA; 33713 MW; 4A7B38D158709B5E CRC64;

Query Match
 Best Local Similarity 76.9%; Score 30; DB 4; Length 311;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
 Db 185 GVNAXSSAF 193

RESULT 12

Q9NP13 PRELIMINARY; PRT; 355 AA.

AC Q9NP13
 DT 01-OCT-2000 (TRENBLREL. 15, Created)
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
 DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
 DE Lysosomal-associated membrane glycoprotein-1 (Fragment).
 GN LAMP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Placenta;
 RX MEDLINE=89066687; PubMed=3198605;
 RA Fukuda M., Vitale J., Matreson J., Carlson S.R.;
 RT "Cloning of cDNAs encoding human lysosomal membrane glycoproteins, h-lamp-1 and h-lamp-2. Comparison of their deduced amino acid

```

RT sequences."
RL J. Biol. Chem. 263:18920-18928(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=93232065; PubMed8517882;
RA Sawada R., Jardine K.A., Fukuda M.;
RT "The genes of major lysosomal membrane glycoproteins, lamp-1 and lamp-
RT 2, 5'-flanking sequence of lamp-2 gene and comparison of exon
RT organization in two genes [published erratum appears in J Biol Chem
RT 1993 Jun 15;268(17):13010].";
RL J. Biol. Chem. 268:9014-9022(1993).
DR EMBL; L08582; AAF6141.1; .
DR EMBL; L08576; AAF6141.1; JOINED.
DR EMBL; L08577; AAF6141.1; JOINED.
DR EMBL; L08578; AAF6141.1; JOINED.
DR EMBL; L08579; AAF6141.1; JOINED.
DR EMBL; L08580; AAF6141.1; JOINED.
DR EMBL; L08581; AAF6141.1; JOINED.
DR InterPro; IP002000; Lamp.
DR Pfam; PF01299; Lamp; 1.
DR PRINTS; PR00336; LYSASOCTDMP.
DR PROSITE; PS00310; LAMP_1; 1.
FT NON_TER
SQ SEQUENCE 355 AA; 38358 MW; 82875ASB384CFAC CRC64;

Query Match
Best Local Similarity 76.9%; Score 30; DB 4; Length 355;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYNAXSSAF 9
Db 230 GYNASSSRF 238

RESULT 13
ID Q8VH34 PRELIMINARY; PRT; 406 AA.
AC Q8VH34;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DR 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DR 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE LAMP-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Ono K., Han J.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY069968; AAL56070.1; -.
DR InterPro; IP002000; Lamp.
DR Pfam; PF01299; Lamp; 1.
DR PRINTS; PR00336; LYSASOCTDMP.
DR PROSITE; PS00310; LAMP_1; 2.
DR PROSITE; PS00311; LAMP_2; 1.
SQ SEQUENCE 406 AA; 43879 MW; C1BD373548ADF85 CRC64;

Query Match
Best Local Similarity 76.9%; Score 30; DB 11; Length 406;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYNAXSSAF 9
Db 280 GYNASSSLF 288

RESULT 14
ID Q9DC13 PRELIMINARY; PRT; 407 AA.
AC Q9DC13;
DT 01-JUN-2001 (TRENBLrel. 17, Created)

```

```

DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Lysosomal membrane glycoprotein 1.
GN LAMP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Sono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barch G.,
RA Blake J., Boffelli D., Bojunga N., Carinci P., de Bonaide M.F.,
RA Brownstein M.J., Butt C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming U.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RL "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
DR EMBL; AK004637; BAB23428.1; -.
DR MGD; MGI:96745; Lamp1.
DR InterPro; IP002000; Lamp.
DR Pfam; PF01299; Lamp; 1.
DR PRINTS; PR00336; LYSASOCTDMP.
DR PROSITE; PS00310; LAMP_1; 2.
DR PROSITE; PS00311; LAMP_2; 1.
SQ SEQUENCE 407 AA; 43536 MW; 696DC97F627DA84 CRC64;

Query Match
Best Local Similarity 76.9%; Score 30; DB 11; Length 407;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYNAXSSAF 9
Db 281 GYNASSSLF 289

RESULT 15
ID Q8WU33 PRELIMINARY; PRT; 417 AA.
AC Q8WU33;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DR 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DR 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021288; AAH21288.1; -.
DR InterPro; IP002000; Lamp.
DR Pfam; PF01299; Lamp; 1.
DR PRINTS; PR00336; LYSASOCTDMP.
DR PROSITE; PS00310; LAMP_1; 2.
DR PROSITE; PS00311; LAMP_2; 1.

```

KV Hypothetical protein.
SQ SEQUENCE 417 AA; 44882 MM; 3E0A28574DD6588 CRC64;

Query Match 76.9%; Score 30; DB 4; Length 417;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
|:|:|
Db 291 GVNAXSSRF 299

Search completed: January 29, 2004, 14:56:51
Job time : 27.8676 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:52:19 : Search time 11.1176 Seconds
(without alignments)
34.252 Million cell updates/sec

Title: US-10-032-950-7

Perfect score: 39

Sequence: 1 GYNAXSSAF 9

Scoring table: ELOSUM62

Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents RA:*

1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	94.9	9	4	US-09-339-511-7
2	32	82.1	9	4	US-09-339-511-1
3	31	79.5	9	4	US-08-861-476C-3
4	31	79.5	47	4	US-08-861-476C-6
5	30	76.9	33	1	US-08-073-807A-17
6	30	76.9	380	1	US-08-073-807A-18
7	30	76.9	4.6	1	US-08-073-807A-2
8	30	76.9	4.6	4	US-09-311-784A-26
9	30	76.9	6.4	4	US-09-832-498-2
10	30	76.9	6.4	4	US-09-832-614A-2
11	29	74.4	9	4	US-09-339-511-5
12	29	74.4	9	4	US-09-339-511-6
13	29	74.4	273	4	US-09-663-600A-103
14	29	74.4	273	4	US-09-663-600A-197
15	28	71.8	9	4	US-09-339-511-2
16	28	71.8	109	4	US-09-199-637A-69
17	28	71.8	264	3	US-09-053-197A-8
18	28	71.8	264	4	US-09-085-761A-8
19	28	71.8	281	4	US-08-040-548-3
20	28	71.8	281	1	US-08-466-344-3
21	28	71.8	284	4	US-09-328-352-6033
22	28	71.8	329	1	US-08-230-047-7
23	28	71.8	4.6	1	US-09-041-075A-2
24	28	71.8	439	4	US-09-328-352-4368
25	28	71.8	436	2	US-08-224-482-2
26	28	71.8	533	1	US-08-040-548-1
27	28	71.8	533	1	US-08-466-344-1

28	28	71.8	533	6	5206152-2	Patent No. 5206152
29	27	69.2	237	3	US-09-111-470-2	Sequence 2, Appli
30	27	69.2	248	3	US-08-750-145A-20	Sequence 20, Appli
31	27	69.2	248	3	US-08-975-638A-24	Sequence 24, Appli
32	27	69.2	248	3	US-09-417-030-24	Sequence 24, Appli
33	27	69.2	248	4	US-09-727-578-24	Sequence 24, Appli
34	27	69.2	319	4	US-09-724-510-2	Sequence 2, Appli
35	27	69.2	319	4	US-09-723-216-2	Sequence 2, Appli
36	27	69.2	319	4	US-09-675-227-2	Sequence 2, Appli
37	27	69.2	343	4	US-09-967-968A-8	Sequence 8, Appli
38	27	69.2	363	4	US-09-967-908A-10	Sequence 10, Appli
39	27	69.2	373	4	US-09-967-908A-4	Sequence 4, Appli
40	27	69.2	381	4	US-09-967-908A-6	Sequence 6, Appli
41	27	69.2	382	1	US-08-186-833-2	Sequence 2, Appli
42	27	69.2	579	4	US-09-521-780-5	Sequence 5, Appli
43	27	69.2	579	4	US-09-521-780-6	Sequence 6, Appli
44	27	69.2	634	4	US-09-328-352-4627	Sequence 4627, Ap
45	27	69.2	758	1	US-08-258-158-2	Sequence 2, Appli

ALIGNMENTS

```

RESULT 1
US-09-339-511-7
Sequence 7, Application US/09339511
Patent No. 6337385
GENERAL INFORMATION:
APPLICANT: Muir, Tom
APPLICANT: Mayville, Patricia
APPLICANT: No. 6337385ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: J1, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
TITLE OF INVENTION: INTERFERENCE
FILE REFERENCE: 600-1-231n
CURRENT APPLICATION NUMBER: US/09/339,511
CURRENT FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-7
Query Match          94.9%: Score 37; DB 4; Length 9;
Best Local Similarity 100.0%: Pred. No. 2.5e+05;
Matches          9; Conservative          0; Mismatches          0; Gaps          0;
OY          1 GYNAXSSAF 9
Db          1 GYNAXSSAF 9
RESULT 2
US-09-339-511-1
Sequence 1, Application US/09339511
Patent No. 6337385
GENERAL INFORMATION:
APPLICANT: Muir, Tom
APPLICANT: Mayville, Patricia
APPLICANT: No. 6337385ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: J1, Guangyong

```

TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
INTERFERENCE
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/09/339,511
CURRENT FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 1
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: peptide
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-1

Query Match
Best Local Similarity 82.1%; Score 32; DB 4; Length 9;
88.9%; Pred. No. 2.5e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
|||
DB 1 GVNAXSSLF 9

RESULT 3
US-08-861-476C-3
Sequence 3, Application US/08861476C
Patent No. 6447786
GENERAL INFORMATION:
APPLICANT: New York University Medical Center
TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
FILE REFERENCE: 63753/7
CURRENT APPLICATION NUMBER: US/08/861,476C
CURRENT FILING DATE: 1997-05-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 3
LENGTH: 9
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-08-861-476C-3

Query Match
Best Local Similarity 79.5%; Score 31; DB 4; Length 9;
77.8%; Pred. No. 2.5e+05;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
|||
DB 1 GVNAXSSLF 9

RESULT 4
US-08-861-476C-6
Sequence 6, Application US/08861476C
Patent No. 6447786
GENERAL INFORMATION:
APPLICANT: New York University Medical Center
TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
FILE REFERENCE: 63753/7
CURRENT APPLICATION NUMBER: US/08/861,476C
CURRENT FILING DATE: 1997-05-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 6
LENGTH: 47
TYPE: PRT
ORGANISM: Staphylococcus aureus

US-08-861-476C-6

Query Match
Best Local Similarity 79.5%; Score 31; DB 4; Length 47;
77.8%; Pred. No. 2.7;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
|||
DB 24 GVNAXSSLF 32

RESULT 5
US-08-073-807A-17
Sequence 17, Application US/08073807A
Patent No. 5646248
GENERAL INFORMATION:
APPLICANT: Sawada, Ritsuko
APPLICANT: Lowe, John B.
TITLE OF INVENTION: CELL SURFACE LAMP EXPRESSION AND
TITLE OF INVENTION: SELECTIN-DEPENDENT ADHESION
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/073,807A
FILING DATE: 08-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9567
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 353 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-073-807A-17

Query Match
Best Local Similarity 76.9%; Score 30; DB 1; Length 353;
66.7%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
|||
DB 263 GVNAXSSRF 271

RESULT 6
US-08-073-807A-18
Sequence 18, Application US/08073807A
Patent No. 5646248
GENERAL INFORMATION:
APPLICANT: Sawada, Ritsuko
APPLICANT: Lowe, John B.
TITLE OF INVENTION: CELL SURFACE LAMP EXPRESSION AND
TITLE OF INVENTION: SELECTIN-DEPENDENT ADHESION
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/073,807A
; FILING DATE: 08-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 9567
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
; US-08-073-807A-18
;
; Query Match
; Best Local Similarity 76.9%; Score 30; DB 1; Length 380;
; Pred. No. 53;
; Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
;
; QY 1 GYMNASSE 9
; DB 290 GYMNASSE 298
;
; RESULT 7
; US-08-073-807A-2
; Sequence 2, Application US/08073807A
; Patent No. 5646248
; GENERAL INFORMATION:
; APPLICANT: Sawada, Ritsuko
; APPLICANT: Lowe, John B.
; APPLICANT: Fukuda, Minoru
; TITLE OF INVENTION: CELL SURFACE LAMP EXPRESSION AND
; TITLE OF INVENTION: SELECTIN-DEPENDENT ADHESION
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/073,807A
; FILING DATE: 08-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 9567
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 416 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-073-807A-2
;
; Query Match
; Best Local Similarity 76.9%; Score 30; DB 1; Length 416;
; Pred. No. 59;
; Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
;
; QY 1 GYMNASSE 9
; DB 290 GYMNASSE 298
;
; RESULT 9
; US-09-311-784A-26
; Sequence 26, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: EpiImmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.01
; CURRENT APPLICATION NUMBER: US/09/311,784A
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human lysosomal membrane glycoprotein-1 (LAMP-1)
;
; US-09-311-784A-26
;
; Query Match
; Best Local Similarity 76.9%; Score 30; DB 4; Length 416;
; Pred. No. 59;
; Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
;
; QY 1 GYMNASSE 9
; DB 290 GYMNASSE 298
;
; RESULT 9
; US-09-832-498-2
; Sequence 2, Application US/09832498
; Patent No. 6410291
; GENERAL INFORMATION:
; APPLICANT: Danielson, Steffen
; APPLICANT: Schneider, Palte
; TITLE OF INVENTION: Polypeptides having haloperoxidase activity
; FILE REFERENCE: 10040,200-US
; CURRENT APPLICATION NUMBER: US/09/832,498
; FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Geniculospirium sp.
;
; US-09-832-498-2

```

Query Match 76.9%; Score 30; DB 4; Length 614;
 Best Local Similarity 66.7%; Pred. No. 94;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNXXSASF 9
 |||:|||||:
 Db 151 GVDASSSF 159

RESULT 10
 US-09-832-614A-2
 ; Sequence 2, Application US/09832614A
 ; Patent No. 6410292
 ; GENERAL INFORMATION:
 ; APPLICANT: No. 6410292zymes A/S
 ; APPLICANT: Danielson, Steffen
 ; APPLICANT: Schneider, Palle
 ; TITLE OF INVENTION: Nucleic acids encoding polypeptides having haloperoxidase activity
 ; FILE REFERENCE: 10173-200-US
 ; CURRENT APPLICATION NUMBER: US/09/832,614A
 ; CURRENT FILING DATE: 2001-04-11
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 614
 ; TYPE: PRT
 ; ORGANISM: Geniculisporium sp.
 ; US-09-832-614A-2

Query Match 76.9%; Score 30; DB 4; Length 614;
 Best Local Similarity 66.7%; Pred. No. 94;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNXXSASF 9
 |||:|||||:
 Db 151 GVDASSSF 159

RESULT 11
 US-09-339-511-5
 ; Sequence 5, Application US/09339511
 ; Patent No. 6337385
 ; GENERAL INFORMATION:
 ; APPLICANT: Milt, Tom
 ; APPLICANT: Mayville, Patricia
 ; APPLICANT: No. 6337385ick, Richard P.
 ; APPLICANT: Beavis, Ronald
 ; APPLICANT: Ji, Guangyong
 ; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
 ; FILE REFERENCE: 600-1-231N
 ; CURRENT APPLICATION NUMBER: US/09/339,511
 ; CURRENT FILING DATE: 1999-06-24
 ; PRIOR APPLICATION NUMBER: 60/090,402
 ; PRIOR FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE: Description of Artificial Sequence: Synthetic
 ; OTHER INFORMATION: peptide
 ; NAME/KEY: VARIANT
 ; LOCATION: (5)
 ; OTHER INFORMATION: Xaa represents any amino acid at this position.
 ; US-09-339-511-5

Query Match 74.4%; Score 29; DB 4; Length 9;
 Best Local Similarity 77.8%; Pred. No. 2.5e+05;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNXXSASF 9
 |||:|||||:
 Db 1 GVNXXSASF 9

RESULT 12
 US-09-339-511-6
 ; Sequence 6, Application US/09339511
 ; Patent No. 6337385
 ; GENERAL INFORMATION:
 ; APPLICANT: Milt, Tom
 ; APPLICANT: Mayville, Patricia
 ; APPLICANT: No. 6337385ick, Richard P.
 ; APPLICANT: Beavis, Ronald
 ; APPLICANT: Ji, Guangyong
 ; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
 ; FILE REFERENCE: 600-1-231N
 ; CURRENT APPLICATION NUMBER: US/09/339,511
 ; CURRENT FILING DATE: 1999-06-24
 ; PRIOR APPLICATION NUMBER: 60/090,402
 ; PRIOR FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 6
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE: Description of Artificial Sequence: Synthetic
 ; OTHER INFORMATION: peptide
 ; NAME/KEY: VARIANT
 ; LOCATION: (5)
 ; OTHER INFORMATION: Xaa represents any amino acid at this position.
 ; US-09-339-511-6

Query Match 74.4%; Score 29; DB 4; Length 9;
 Best Local Similarity 77.8%; Pred. No. 2.5e+05;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNXXSASF 9
 |||:|||||:
 Db 1 GVNXXSASF 9

RESULT 13
 US-09-663-600A-103
 ; Sequence 103, Application US/09663600A
 ; Patent No. 6573068
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
 ; APPLICANT: Duclelet, Aymeric
 ; APPLICANT: Bougueleret, Lydie
 ; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
 ; FILE REFERENCE: 31. US3. CIP
 ; CURRENT APPLICATION NUMBER: US/09/663,600A
 ; CURRENT FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: 09/191,997
 ; PRIOR FILING DATE: 1998-11-13
 ; PRIOR APPLICATION NUMBER: 60/066,677
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/069,957
 ; PRIOR FILING DATE: 1997-12-17
 ; PRIOR APPLICATION NUMBER: 60/074,121
 ; PRIOR FILING DATE: 1998-02-09
 ; PRIOR APPLICATION NUMBER: 60/081,563
 ; PRIOR FILING DATE: 1998-04-13
 ; PRIOR APPLICATION NUMBER: 60/096,116
 ; PRIOR FILING DATE: 1998-08-10
 ; PRIOR APPLICATION NUMBER: 60/099,273
 ; PRIOR FILING DATE: 1998-09-04
 ; NUMBER OF SEQ ID NOS: 229
 ; SOFTWARE: PatentIn Ver. 2.0


```

; SEQ ID NO 103
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
;   NAME/KEY: SIGNAL
;   LOCATION: -45...-1
;   NAME/KEY: UNSURE
;   LOCATION: 181,187,193,196,198,199,203,212,214
; OTHER INFORMATION: Xaa = any one of the twenty amino acids
US-09-663-600A-103

```

```

Query Match          74.4%; Score 29; DB 4; Length 273;
Best Local Similarity 66.7%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 GVNAXSSAF 9
Db      12 GVNKYSTAF 20

```

```

RESULT 14
US-09-663-600A-197
; Sequence 197, Application US/09663600A
; Patent No. 6573068
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bouquelaret, Lydie
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: 31, US3 CIP
; CURRENT APPLICATION NUMBER: US/09/663,600A
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent.pm
; SEQ ID NO 197
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
;   NAME/KEY: SIGNAL
;   LOCATION: -45...-1
US-09-663-600A-197

```

```

Query Match          74.4%; Score 29; DB 4; Length 273;
Best Local Similarity 66.7%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 GVNAXSSAF 9
Db      12 GVNKYSTAF 20

```

```

RESULT 15
US-09-339-511-2
; Sequence 2, Application US/09339511
; Patent No. 6337385
; GENERAL INFORMATION:
; APPLICANT: Mullr, Tom

```

```

; APPLICANT: Mayville, Patricia
; APPLICANT: No. 6337385ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-221N
; CURRENT APPLICATION NUMBER: US/09/339,511
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentm Ver. 2.0
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-2

```

```

Query Match          71.8%; Score 28; DB 4; Length 9;
Best Local Similarity 77.8%; Pred. No. 2.5e+05;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 GVNAXSSAF 9
Db      1 GANAXSLF 9

```

```

Search completed: January 29, 2004, 15:03:22
Job time : 11.1176 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: January 29, 2004, 14:47:39 ; Search time 25.5441 Seconds
(without alignments)
73.223 Million cell updates/sec

Title: US-10-032-950-7
Perfect score: 39
Sequence: 1 GYNASSAF 9

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications AA:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	37	94.9	9 13 US-10-032-950-7	Sequence 7, Appli
2	32	82.1	9 13 US-10-032-950-1	Sequence 1, Appli
3	31	79.5	9 15 US-10-201-444-3	Sequence 3, Appli
4	31	79.5	47 15 US-10-201-444-6	Sequence 6, Appli
5	30	76.9	163 16 US-10-080-170-155	Sequence 155, App
6	30	76.9	167 16 US-10-080-170-538	Sequence 538, App
7	30	76.9	416 12 US-10-371-669-26	Sequence 26, Appli
8	30	76.9	416 12 US-10-371-665-26	Sequence 26, Appli
9	30	76.9	416 12 US-10-371-657-2	Sequence 2, Appli
10	30	76.9	614 14 US-10-151-557-2	Sequence 2, Appli
11	29	74.4	9 13 US-10-032-950-5	Sequence 5, Appli
12	29	74.4	9 13 US-10-032-950-6	Sequence 6, Appli
13	29	74.4	273 12 US-10-319-763-103	Sequence 103, App
14	29	74.4	273 12 US-10-319-763-197	Sequence 197, App
15	29	74.4	273 12 US-09-981-151A-64	Sequence 64, Appli

16	29	74.4	273 12 US-09-981-151A-65	Sequence 65, Appli
17	29	74.4	273 12 US-10-295-027-344	Sequence 344, App
18	29	74.4	273 12 US-10-295-027-1550	Sequence 1250, Ap
19	29	74.4	287 15 US-10-106-698-4775	Sequence 4775, Ap
20	29	74.4	3635 10 US-09-845-589-2	Sequence 2, Appli
21	29	74.4	3635 10 US-10-037-182-4	Sequence 4, Appli
22	28	71.8	9 13 US-10-032-950-2	Sequence 2, Appli
23	28	71.8	80 9 US-09-864-761-43395	Sequence 89, Appli
24	28	71.8	109 11 US-09-975-719-89	Sequence 10071, A
25	28	71.8	405 15 US-10-156-761-10071	Sequence 2, Appli
26	28	71.8	416 9 US-09-742-582-2	Sequence 2, Appli
27	28	71.8	416 11 US-09-742-580-2	Sequence 2, Appli
28	28	71.8	416 11 US-09-742-581-2	Sequence 12518, A
29	28	71.8	451 15 US-10-156-761-1518	Sequence 17, Appli
30	28	71.8	533 12 US-10-373-238-17	Sequence 38, Appli
31	28	71.8	583 9 US-09-810-264-38	Sequence 183, App
32	28	71.8	998 12 US-10-116-275-183	Sequence 1488, Ap
33	27	69.2	55 12 US-09-864-408A-1488	Sequence 8250, Ap
34	27	69.2	67 15 US-10-156-761-8250	Sequence 600, App
35	27	69.2	144 12 US-09-864-408A-600	Sequence 2, Appli
36	27	69.2	237 10 US-09-862-802-2	Sequence 49, Appli
37	27	69.2	237 10 US-09-870-759-49	Sequence 2653, Ap
38	27	69.2	237 12 US-09-751-708A-49	Sequence 464, App
39	27	69.2	237 12 US-10-264-237-2653	Sequence 303, App
40	27	69.2	243 9 US-09-764-870-464	Sequence 303, App
41	27	69.2	243 15 US-10-125-540-464	Sequence 1925, A
42	27	69.2	246 5 US-09-764-870-303	Sequence 727, App
43	27	69.2	247 12 US-10-125-540-303	
44	27	69.2	247 12 US-10-369-493-1925	
45	27	69.2	269 12 US-10-369-493-727	

ALIGNMENTS

RESULT 1
US-10-032-950-7
Sequence 7, Application US/10032950
Publication No. US2002007453A1
GENERAL INFORMATION:
APPLICANT: Mult, Tom
APPLICANT: Mayville, Patricia
APPLICANT: No. US2002007453A1ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: Ji, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
TITLE OF INVENTION: INTERFERENCE
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/10/032,950
PRIOR APPLICATION NUMBER: 2001-12-27
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 7
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-7

Query Match
Best Local Similarity 94.9% Score 37; DB 13; Length 9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 GYNASSAF 9
1 GYNASSAF 9

RESULT 2

US-10-032-950-1
; Sequence 1, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Muir, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; TITLE OF INVENTION: INTERFERENCE
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-1

Query Match 82.1%; Score 32; DB 13; Length 9;
Best Local Similarity 88.9%; Pred. No. 7e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9

Db 1 GVNAXSSLF 9

RESULT 3

US-10-201-444-3
; Sequence 3, Application US/10201444
; Publication No. US20030076378A1
; GENERAL INFORMATION:
; APPLICANT: New York University Medical Center
; TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
; FILE REFERENCE: 63753/7
; CURRENT APPLICATION NUMBER: US/10/201,444
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/08/861,476
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-201-444-3

Query Match 79.5%; Score 31; DB 15; Length 9;
Best Local Similarity 77.8%; Pred. No. 7e+05;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9

Db 1 GVNAXSSLF 9

RESULT 4

US-10-201-444-6

; Sequence 6, Application US/10201444
; Publication No. US20030076378A1
; GENERAL INFORMATION:
; APPLICANT: New York University Medical Center
; TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
; FILE REFERENCE: 63753/7
; CURRENT APPLICATION NUMBER: US/10/201,444
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US/08/861,476
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-201-444-6

Query Match 79.5%; Score 31; DB 15; Length 47;
Best Local Similarity 77.8%; Pred. No. 6.8e+05;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9

Db 24 GVNAXSSLF 32

RESULT 5

US-10-080-170-155
; Sequence 155, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 155
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-080-170-155

Query Match 76.9%; Score 30; DB 16; Length 163;
Best Local Similarity 77.8%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9

Db 132 GVNAXSSAF 140

RESULT 6

US-10-080-170-538
; Sequence 538, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 538
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-538
```

```
Query Match          76.9%; Score 30; DB 16; Length 167;
Best Local Similarity 77.8%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 GVNAXSSAF 9
        |||||
Db       136 GVNAXSSAF 144
```

```
RESULT 7
US-10-371-069-26
; Sequence 26, Application US/10371069
; Publication No. US20030216342A1
; GENERAL INFORMATION:
```

```
; APPLICANT: EPIMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
```

```
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE REFERENCE: 39963-20022.10
; CURRENT APPLICATION NUMBER: US/10/371,069
; PRIOR FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; OTHER INFORMATION: human lysosomal membrane glycoprotein-1 (LAMP-1)
US-10-371-069-26
```

```
Query Match          76.9%; Score 30; DB 12; Length 416;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 GVNAXSSAF 9
        |||||
Db       290 GVNAXSSAF 298
```

```
RESULT 8
US-10-371-645-26
; Sequence 26, Application US/10371645
; Publication No. US20030216343A1
; GENERAL INFORMATION:
```

```
; APPLICANT: EPIMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE REFERENCE: 39963-20022.11
```

```
; CURRENT APPLICATION NUMBER: US/10/371,645
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; OTHER INFORMATION: human lysosomal membrane glycoprotein-1 (LAMP-1)
US-10-371-645-26
```

```
Query Match          76.9%; Score 30; DB 12; Length 416;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 GVNAXSSAF 9
        |||||
Db       290 GVNAXSSAF 298
```

```
RESULT 9
US-10-371-260-26
; Sequence 26, Application US/10371260
; Publication No. US20030220285A1
; GENERAL INFORMATION:
```

```
; APPLICANT: EPIMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
```

```
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE REFERENCE: 39963-20022.13
; CURRENT APPLICATION NUMBER: US/10/371,260
; PRIOR FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; OTHER INFORMATION: human lysosomal membrane glycoprotein-1 (LAMP-1)
US-10-371-260-26
```

```
Query Match          76.9%; Score 30; DB 12; Length 416;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 GVNAXSSAF 9
        |||||
Db       290 GVNAXSSAF 298
```

```
RESULT 10
US-10-151-557-2
; Sequence 2, Application US/10151557
; Publication No. US20020183506A1
; GENERAL INFORMATION:
```

```
; APPLICANT: No. US20020183506A1ozymes A/S
; APPLICANT: Danielsen, Steffen
```

```

; APPLICANT: Schneider, Palle
; TITLE OF INVENTION: Nucleic acids encoding polypeptides having
; FILE REFERENCE: 10173.200-US
; CURRENT APPLICATION NUMBER: US/10/151,557
; PRIOR APPLICATION NUMBER: US/09/832,614A
; PRIOR FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Geniculosporium sp.
; US-10-151-557-2

```

```

Query Match      76.9%; Score 30; DB 14; Length 614;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 GVNAXSSAF 9
Db      151 GVDASSSF 159

```

```

RESULT 11
US-10-032-950-5
; Sequence 5, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mullr, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beaver, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
; US-10-032-950-5

```

```

Query Match      74.4%; Score 29; DB 13; Length 9;
Best Local Similarity 77.8%; Pred. No. 7e+05;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 GVNAXSSAF 9
Db      1 GVNAXSSAF 9

```

```

RESULT 12
US-10-032-950-6
; Sequence 6, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mullr, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.

```

```

; APPLICANT: Beaver, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
; US-10-032-950-6

```

```

Query Match      74.4%; Score 29; DB 13; Length 9;
Best Local Similarity 77.8%; Pred. No. 7e+05;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 GVNAXSSAF 9
Db      1 GVNAXSSAF 9

```

```

RESULT 13
US-10-319-763-103
; Sequence 103, Application US/10319763
; Publication No. US20030144430A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: G-031.US04.DIV
; CURRENT APPLICATION NUMBER: US/10/319,763
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent. pm
; SEQ ID NO 103
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -45...-1
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 181,187,193,196,198,199,203,212,214
; OTHER INFORMATION: Xaa = any one of the twenty amino acids
; US-10-319-763-103

```

```

Query Match      74.4%; Score 29; DB 12; Length 273;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;

```

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
 ||| |:
 Db 12 GVNKYSTAF 20

RESULT 14

US-10-319-763-197
 ; Sequence 197, Application US/10319763
 ; Publication No. US2003014490A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Mline Edwards, Jean-Baptiste
 ; APPLICANT: Duclert, Aymeric
 ; APPLICANT: Bougueleret, Lydie
 ; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
 ; FILE REFERENCE: G-031.US04.DIV
 ; CURRENT APPLICATION NUMBER: US/10/319,763
 ; CURRENT FILING DATE: 2002-12-10
 ; PRIOR APPLICATION NUMBER: 60/066,677
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/069,957
 ; PRIOR FILING DATE: 1997-12-17
 ; PRIOR APPLICATION NUMBER: 60/074,121
 ; PRIOR FILING DATE: 1998-02-09
 ; PRIOR APPLICATION NUMBER: 60/081,563
 ; PRIOR FILING DATE: 1998-04-13
 ; PRIOR APPLICATION NUMBER: 60/096,116
 ; PRIOR FILING DATE: 1998-08-10
 ; PRIOR APPLICATION NUMBER: 60/099,273
 ; PRIOR FILING DATE: 1998-09-04
 ; NUMBER OF SEQ ID NOS: 229
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 197
 ; LENGTH: 273
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SIGNAL
 ; LOCATION: -45...-1
 ; US-10-319-763-197

Query Match 74.4%; Score 29; DB 12; Length 273;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
 ||| |:
 Db 12 GVNKYSTAF 20

RESULT 15

US-09-981-151A-64
 ; Sequence 64, Application US/09981151A
 ; Publication No. US20030212256A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Edinger, Shlomit R
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: MacDougall, John R
 ; APPLICANT: Malvankar, Muriel M
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Miller, Isabelle
 ; APPLICANT: Peyman, John A
 ; APPLICANT: Stone, David J
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Shinkets, Richard A
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Guo, Xiaojia
 ; APPLICANT: Patuturajan, Weera
 ; APPLICANT: Taupier Jr, Raymond J
 ; APPLICANT: Burgess, Catherine E
 ; APPLICANT: Zernhausen, Bryan D

APPLICANT: Kerkuda, Ramesh
 ; APPLICANT: Spytek, Kimberly A
 ; APPLICANT: Gangolli, Esra A
 ; APPLICANT: Fernandes, Eima R
 ; APPLICANT: Gorman, Linda
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-168
 ; CURRENT APPLICATION NUMBER: US/09/981,151A
 ; CURRENT FILING DATE: 2001-10-16
 ; PRIOR APPLICATION NUMBER: 60/241,040
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/241,058
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/241,063
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/241,243
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/242,152
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/242,482
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/242,611
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/242,612
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/242,880
 ; PRIOR FILING DATE: 2000-10-24
 ; PRIOR APPLICATION NUMBER: 60/242,881
 ; PRIOR FILING DATE: 2000-10-24
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 160
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 64
 ; LENGTH: 273
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-981-151A-64

Query Match 74.4%; Score 29; DB 12; Length 273;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GVNAXSSAF 9
 ||| |:
 Db 12 GVNKYSTAF 20

Search completed: January 29, 2004, 15:01:46
 Job time : 25.541 secs

XX PA (UYNO-) UNIV NOTTINGHAM.
 XX PI Affas ZM, Bycroft BW, Chan WC, McDowell PM, Stewart GSAB,
 PI Williams P;
 XX DR WPI; 1999-370865/31.
 XX PT Cyclic antibacterial peptides
 PS Claim 1; Page -; 32pp; English.
 XX CC The present sequence represents a cyclic peptide which is a modulator
 CC of the agr response. Agonists of this response antagonize early stages
 CC of virulence, specifically production of cell wall proteins (so are
 CC preferred for prophylaxis), while antagonists interfere with the later
 CC stages of virulence, specifically secretion of exotoxin (so are
 CC preferred for treatment). The peptides are used to treat or prevent
 CC bacterial infections in humans and animals, also for incorporation into
 CC medical articles, e.g. tampons (to prevent toxic shock syndrome),
 CC prostheses and wound dressings. Particularly they are active against
 CC staphylococci, specifically Staphylococcus aureus but also
 CC coagulase-negative S. epidermidis, implicated in infection of medical
 CC implants.
 CC note: this sequence does not appear in the specification; it was created
 CC using information provided.
 XX SQ Sequence 4 AA;
 Query Match 94.7%; Score 18; DB 20; Length 4;
 Best Local Similarity 100.0%; Pred. No. 9.3e-05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 SSLF 5
 DB 1 SSLF 4
 RESULT 2
 ID AA67858 standard; peptide; 5 AA.
 XX AC AA67858;
 XX DT 25-APR-2000 (first entry)
 XX DE S. aureus peptide #6 used for bacterial interference.
 XX KM Staphylococcus aureus infection; cyclic peptide; AgrD; agr response;
 KM virulence factor; treatment.
 XX OS Staphylococcus aureus.
 XX FH Key Location/Qualifiers
 FT Misc-difference 1 /label= Unknown
 FT /note= "N-terminal residue is linked to the C-terminal
 residue to form a cyclic peptide"
 XX WO9967286-A2.
 XX PD 29-DEC-1999.
 XX PF 24-JUN-1999; 99WO-US14562.
 XX PR 24-JUN-1998; 98US-0103438.
 XX PA (UYRO) UNIV ROCKEFELLER.
 XX PA (UYNO) UNIV NEW YORK STATE.
 XX PI Muir TW, Mayville P, Novick RP, Ji G, Beavis R;
 XX DR WPI; 2000-147202/13.

XX PT New cyclic peptides for treating infections with Staphylococcus aureus
 PT
 XX PS Claim 9; Page 26; 37pp; English.
 XX CC This sequence represents a cyclic peptide derived from the Staphylococcus
 CC aureus AgrD peptide. The invention relates to AgrD derived peptides,
 CC a composition containing a peptide and a carrier, and a method for the
 CC production of the cyclic peptides. The peptide inhibits the agr response,
 CC which is normally associated with the release of virulence factors of
 CC Staphylococcus aureus. An AgrD peptide is produced by S. aureus that
 CC activates the agr response in strains of a single group, but interferes
 CC with this response in strains of different groups. The peptides and
 CC composition containing them can be used to treat infections by S. aureus.
 XX SQ Sequence 5 AA;
 Query Match 94.7%; Score 18; DB 21; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e-05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 SSLF 5
 DB 2 SSLF 5
 RESULT 3
 ID ABP53547 standard; peptide; 5 AA.
 XX AC ABP53547;
 XX DT 13-DEC-2002 (first entry)
 XX DE Cyclic peptide SEQ ID NO:8.
 XX KM Cyclic peptide; Staphylococcus aureus; infection; antibacterial;
 XX agr response inhibitor.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "any amino acid"
 XX US2002077453-A1.
 XX PD 20-JUN-2002.
 XX PF 27-DEC-2001; 2001US-0032950.
 XX PR 24-JUN-1998; 98US-090402P.
 XX PR 24-JUN-1999; 99US-0339511.
 XX PA (MUIR/) MUIR T W.
 PA (MAYV/) MAYVILLE P.
 PA (NOVI/) NOVICK R P.
 PA (BEAV/) BEAVIS R.
 PA (JIGG/) JI G.
 XX PI Muir TW, Mayville P, Novick RP, Beavis R, Ji G;
 XX DR WPI; 2002-681366/73.
 XX PT New cyclic peptides, useful for treating Staphylococcus aureus
 PT infections
 XX PS Claim 9; Page 10; 18pp; English.
 CC ABP53540 to ABP53547 represent cyclic peptides (I) from the present
 CC invention. The present invention also describes a method for treating
 CC Staphylococcus aureus infection comprising the administration of a

CC composition comprising (1). (1) has antibacterial activity, and can be
 CC used as an agr gene response inhibitor. The peptides are useful for
 CC treating *S. aureus* infections.

XX Sequence 5 AA;

Query Match 94.7%; Score 18; DB 23; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSUF 5

Db 2 SSUF 5

RESULT 4
 AAM50906
 ID AAM50906 standard; Peptide; 5 AA.

XX AAM50906;

AC 08-MAY-2002 (first entry)

XX AgrD-autoinducing cyclic peptide, inhibitor of agr response.

XX Staphylococcus aureus; AgrD; agr response: inhibitor; antibiotic;
 XX antibacterial; infection; therapy; cyclic.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "any amino acid"

FT Misc-difference 1 /note= "note linked to residue 5 to form cyclic
 peptide"

FT Misc-difference 5 /note= "note linked to residue 1 to form cyclic
 peptide"

XX US6337385-B1.

XX 08-JAN-2002.

XX 24-JUN-1999; 99US-0339511.

XX 24-JUN-1998; 98US-090402P.

XX (TYRO) UNIV ROCKEFELLER.

XX (UYN) UNIV NEW YORK STATE.

XX Milr TW, Mayville P, Novick RP, Beavis R, U G;

XX WPI; 2002-170774/22.

XX Novel synthetic, cyclic AgrD-autoinducing peptide for bacterial
 XX interference and for treating *Staphylococcus aureus* infection in a
 XX subject

XX Claim 7; Column 19; 18pp; English.

XX The present sequence is that of a novel synthetic cyclic peptide
 CC of the invention that is capable of inhibiting the agr response of
 CC *Staphylococcus aureus*. It is an AgrD-autoinducing peptide, where
 CC AgrD is a secreted agr-encoded peptide and where the agr locus
 CC controls the synthesis of virulence factor and other extracellular
 CC proteins responsible for pathogenicity in *S. aureus*. Preferred
 CC peptides may have the sequence NH₂-X(n)-Z-X(y)-COOH, with a
 CC cyclic bond between the Z residue and COOH other than a thioester
 CC bond, where X is an amino acid, an amino acid analogue, a
 CC peptidomimetic or non-amide isostere, Z is a synthetic or a
 CC biosynthetic amino acid, n is 0-10 and y is 1-10. The cyclic bond
 CC is especially a lactam or lactone bond. The thiololactone

CC structure within native AgrD peptides is required for activation
 CC of the agr response. Elimination of the thiol ester component of
 CC the cyclic ring structure can destroy agr response activating
 CC activity while preserving and enhancing inhibitory activity. A
 CC claimed method of preparing a cyclic peptide involves: assembling
 CC a linear peptide chain on to a solid phase resin support;
 CC deprotecting the resulting protected assembled peptide; treating the
 CC deprotected peptide with neutral buffer for a time sufficient to
 CC form the cyclic peptide and cleave the peptide from the support;
 CC and recovering the cyclic peptide. The peptide is useful for
 CC bacterial interference, especially for the treatment of *S. aureus*
 CC infection.

XX Sequence 5 AA;

Query Match 94.7%; Score 18; DB 23; Length 5;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSUF 5

Db 2 SSUF 5

RESULT 5
 AAR10519
 ID AAR10519 standard; Protein; 6 AA.

XX AAR10519;

XX 15-APR-1991 (first entry)

XX Fusinus inhibitory peptide B1.

XX muscular contraction-inhibitory neuropeptide; F1Pb1; Mollusca;
 XX radular tricle muscle.

XX Fusinus perplep ferrugineus.

XX JP02306995-A.

XX 20-DEC-1990.

XX 19-MAY-1989; 89JP-0125596.

XX 19-MAY-1989; 89JP-0125596.

XX (SUNR) SUNTORY LTD.

XX WPI; 1991-040149/06.

XX Muscular contraction-inhibitory neuro-peptide(s) in mollusca -
 XX prepd. by purifying oligopeptide using high power lig.
 XX PT chromatography after extracting neuro:ganglion of Fusinus perplep
 XX ferrugineus

XX Claim 6; Page 1; 10pp; Japanese.

XX The peptide was isolated from Fusinus perplep ferrugineus
 CC neuroganga and inhibits muscular contraction, especially of the
 CC radular tricle muscle in *F. perplep ferrugineus*, Prosobranchia
 CC and other Mollusca. Alternatively, it can be synthesised by solid-
 CC phase techniques. One or both of the Ser residues at positions 1 or
 CC 2 may be absent. The peptide can be used in biological studies for
 CC the development of drugs. See also AAR10516-R10518 and AAR10520.

XX Sequence 6 AA;

Query Match 94.7%; Score 18; DB 12; Length 6;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSUF 5

Db 1 SSLF 4

RESULT 6
AAW10783 standard; peptide; 6 AA.

AAW10783;
25-MAR-2003 (updated)
10-NOV-1997 (first entry)

Ferritin motif #26 important for selective binding affinity.

Functional surrogate; analyte; affinity receptor; immunoreactive group; mimic; homogeneous immunoassay; detection; diagnostic analyte; Chlamydia; cardiac marker; tumour marker; allergen; hormone; fertility; myoglobin; pregnancy; infectious disease; ferritin; myosin light chain; tropomyosin; follicle stimulating hormone; human; growth hormone; immunoglobulin E; prolactin; parathyroid hormone; placental lactogen; hepatitis antigen; antibody; chorionic gonadotropin; insulinising hormone; cytomegalovirus; Streptococcus; rubella; toxoplasma; DK-MB; prostate-specific antigen; carcinoembryonic antigen; alpha-fetoprotein; herpes virus; CA125.

Synthetic.
WO9641172-A1.
19-DEC-1996.
07-JUN-1996; 96WO-US10498.
07-JUN-1995; 95US-0476375.
(CYTO-) CYTOGEN CORP.
Lee-owen FV, Carter JM;
WPI; 1997-077284/07.

Labelled functional surrogate of an analyte - useful as competitor molecule in affinity assays, esp. for detecting large macromolecules such as ferritin

Claim 51; Page 56; 156pp; English.

This sequence represents a peptide motif derived from ferritin which is important for selective binding affinity. Peptides containing motifs such as this may be used as functional surrogates in the conjugate of the invention. The novel labelled conjugate comprises at least one CC label attached to a functional surrogate of an analyte of interest. CC The surrogate is capable of competing effectively with the analyte for a limiting amount of an affinity receptor for the analyte. The CC conjugate exhibits an activity that is altered upon interaction with CC the affinity receptor and this activity can be measured and related to CC the amount of the analyte present in a sample. Functional surrogates CC such as this have an immunoreactive group that allows the surrogate to CC compete effectively and with the analyte for a limiting amount of its CC affinity receptor. Functional surrogates are able to mimic naturally CC occurring analytes. They can be labelled for use in standard competitive CC affinity assays (esp. homogeneous immunoassays) for detecting large CC macromolecules such as polypeptides, polysaccharides, polynucleotides, CC glycoproteins and lipid-containing macromolecules, as well as small CC haptens. Typical diagnostic analytes for detection include cardiac or CC tumour markers, allergens, hormones related to fertility-pregnancy or CC analytes associated with infectious disease. In particular, the assays CC are useful for detecting ferritin, follicle stimulating hormone, human CC growth hormone, immunoglobulin E, prolactin, parathyroid hormone, human CC placental lactogen, hepatitis antigen or antibodies against them, human CC chorionic gonadotropin, human insulinising hormone, cytomegalovirus, CC Chlamydia, Streptococcus A, rubella, toxoplasma, herpes virus, DK-MB, CC myoglobin, myosin light chain, tropomyosin, carcinoembryonic antigen.

CC alpha-fetoprotein, prostate-specific antigen and CA125 (a tumour marker).
CC (Updated on 25-MAR-2003 to correct PI field.)
CC
XX
SQ Sequence 6 AA;
Query Match 94.7%; Score 18; DB 18; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 SSLF 5
Db 3 SSLF 6

RESULT 7
AAW25549 standard; peptide; 6 AA.

AAW25549;
25-MAR-2003 (updated)
10-NOV-1997 (first entry)

Synthetic ferritin peptide #22.

Functional surrogate; analyte; affinity receptor; immunoreactive group; mimic; homogeneous immunoassay; detection; diagnostic analyte; Chlamydia; cardiac marker; tumour marker; allergen; hormone; fertility; myoglobin; pregnancy; infectious disease; ferritin; myosin light chain; tropomyosin; follicle stimulating hormone; human; growth hormone; immunoglobulin E; prolactin; parathyroid hormone; placental lactogen; hepatitis antigen; antibody; chorionic gonadotropin; insulinising hormone; cytomegalovirus; Streptococcus; rubella; toxoplasma; DK-MB; prostate-specific antigen; carcinoembryonic antigen; alpha-fetoprotein; herpes virus; CA125.

Synthetic.
WO9641172-A1.
19-DEC-1996.
07-JUN-1996; 96WO-US10498.
07-JUN-1995; 95US-0476375.
(CYTO-) CYTOGEN CORP.
Lee-owen FV, Carter JM;
WPI; 1997-077284/07.

Labelled functional surrogate of an analyte - useful as competitor molecule in affinity assays, esp. for detecting large macromolecules such as ferritin

Claim 51; Page 26; 156pp; English.

This sequence represents a synthetic ferritin peptide which was used as a functional surrogate in the conjugate of the invention. The novel CC labelled conjugate comprises at least one label attached to a functional CC surrogate of an analyte of interest. The surrogate is capable of CC competing effectively with the analyte for a limiting amount of an CC affinity receptor for the analyte. The conjugate exhibits an activity CC that is altered upon interaction with the affinity receptor and this CC activity can be measured and related to the amount of the analyte present CC in a sample. Functional surrogates such as this have an immunoreactive CC group that allows the surrogate to compete effectively and with the CC analyte for a limiting amount of its affinity receptor. Functional CC surrogates are able to mimic naturally occurring analytes. They can CC be labelled for use in standard competitive affinity assays (esp. CC homogeneous immunoassays) for detecting large macromolecules such as CC polypeptides, polysaccharides, polynucleotides, glycoproteins and CC lipid-containing macromolecules, as well as small haptens. Typical

CC diagnostic analytes for detection include cardiac or tumour markers,
 CC allergens, hormones related to fertility-pregnancy or analytes
 CC associated with infectious disease. In particular, the assays are
 CC useful for detecting ferritin, follicle stimulating hormone, human
 CC growth hormone, immunoglobulin E, prolactin, parathyroid hormone, human
 CC placental lactogen, hepatitis antigens or antibodies against them, human
 CC chorionic gonadotropin, human luteinising hormone, cytomegalovirus,
 CC Chlamydia, streptococcus a, rubella, toxoplasma, herpes virus, DK-MB,
 CC myoglobin, myosin light chain, troponin, carcinoembryonic antigen,
 CC alpha-fetoprotein, prostate-specific antigen and CA125 (a tumour marker).
 CC (updated on 25-MAR-2003 to correct PI field.)
 CC
 XX Sequence 6 AA;
 SQ
 Query Match 94.7%; Score 18; DB 18; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 SSUF 5
 Db 1 SSUF 4
 RESULT 8
 AAM25550
 ID AAM25550 standard; peptide; 6 AA.
 XX
 AC AAM25550;
 XX
 DT 25-MAR-2003 (updated)
 DT 10-NOV-1997 (first entry)
 XX
 DE Synthetic ferritin peptide #23.
 XX
 KW Functional surrogate; analyte; affinity receptor; immunoreactive group;
 KW mimic; homogeneous immunoassay; detection; diagnostic analyte; Chlamydia;
 KW cardiac marker; tumour marker; allergen; hormone; fertility; myoglobin;
 KW pregnancy; infectious disease; ferritin; myosin light chain; troponin;
 KW follicle stimulating hormone; human; growth hormone; immunoglobulin E;
 KW prolactin; parathyroid hormone; placental lactogen; hepatitis antigen;
 KW antibody; chorionic gonadotropin; luteinising hormone; cytomegalovirus;
 KW streptococcus; rubella; toxoplasma; DK-MB; prostate-specific antigen;
 KW carcinoembryonic antigen; alpha-fetoprotein; herpes virus; CA125.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "Biotinylated Glu"
 XX
 PN WO9641172-A1.
 XX
 PD 19-DEC-1996.
 XX
 PF 07-JUN-1996; 96WO-US10498.
 XX
 PR 07-JUN-1995; 95US-0476375.
 XX
 PA (CYTO-) CYTOGEN CORP.
 XX
 PI Lee-owen FV, Carter JM;
 XX
 DR WPI; 1997-077284/07.
 XX
 PT Labelled functional surrogate of an analyte - useful as competitor
 PT molecule in affinity assays, esp. for detecting large macromolecules
 PT such as ferritin
 XX
 PS Claim 51; Page 26; 156bp; English.
 XX
 CC This sequence represents a synthetic ferritin peptide which was used as
 CC a functional surrogate in the conjugate of the invention. The novel
 CC labelled conjugate comprises at least one label attached to a functional

CC surrogate of an analyte of interest. The surrogate is capable of
 CC competing effectively with the analyte for a limiting amount of an
 CC affinity receptor for the analyte. The conjugate exhibits an activity
 CC that is altered upon interaction with the affinity receptor and this
 CC activity can be measured and related to the amount of the analyte present
 CC in a sample. Functional surrogates such as this have an immunoreactive
 CC group that allows the surrogate to compete effectively and with the
 CC analyte for a limiting amount of its affinity receptor. Functional
 CC surrogates are able to mimic naturally occurring analytes (esp.
 CC be labelled for use in standard competitive affinity assays (esp.
 CC for homogeneous immunoassays) for detecting large macromolecules such as
 CC polypeptides, polysaccharides, polynucleotides, glycoproteins and
 CC lipid-containing macromolecules, as well as small haptens. Typical
 CC diagnostic analytes for detection include cardiac or tumour markers,
 CC allergens, hormones related to fertility-pregnancy or analytes
 CC associated with infectious disease. In particular, the assays are
 CC useful for detecting ferritin, follicle stimulating hormone, human
 CC growth hormone, immunoglobulin E, prolactin, parathyroid hormone, human
 CC placental lactogen, hepatitis antigens or antibodies against them, human
 CC chorionic gonadotropin, human luteinising hormone, cytomegalovirus,
 CC Chlamydia, streptococcus a, rubella, toxoplasma, herpes virus, DK-MB,
 CC myoglobin, myosin light chain, troponin, carcinoembryonic antigen,
 CC alpha-fetoprotein, prostate-specific antigen and CA125 (a tumour marker).
 CC (updated on 25-MAR-2003 to correct PI field.)
 CC
 XX Sequence 6 AA;
 SQ
 Query Match 94.7%; Score 18; DB 18; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 SSUF 5
 Db 1 SSUF 4
 RESULT 9
 AAM25555
 ID AAM25555 standard; peptide; 6 AA.
 XX
 AC AAM25555;
 XX
 DT 25-MAR-2003 (updated)
 DT 10-NOV-1997 (first entry)
 XX
 DE Synthetic ferritin peptide #28.
 XX
 KW Functional surrogate; analyte; affinity receptor; immunoreactive group;
 KW mimic; homogeneous immunoassay; detection; diagnostic analyte; Chlamydia;
 KW cardiac marker; tumour marker; allergen; hormone; fertility; myoglobin;
 KW pregnancy; infectious disease; ferritin; myosin light chain; troponin;
 KW follicle stimulating hormone; human; growth hormone; immunoglobulin E;
 KW prolactin; parathyroid hormone; placental lactogen; hepatitis antigen;
 KW antibody; chorionic gonadotropin; luteinising hormone; cytomegalovirus;
 KW streptococcus; rubella; toxoplasma; DK-MB; prostate-specific antigen;
 KW carcinoembryonic antigen; alpha-fetoprotein; herpes virus; CA125.
 XX
 OS Synthetic.
 XX
 PN WO9641172-A1.
 XX
 PD 19-DEC-1996.
 XX
 PF 07-JUN-1996; 96WO-US10498.
 XX
 PR 07-JUN-1995; 95US-0476375.
 XX
 PA (CYTO-) CYTOGEN CORP.
 XX
 PI Lee-owen FV, Carter JM;
 XX
 DR WPI; 1997-077284/07.
 XX

PT Labelled functional surrogate of an analyte - useful as competitor molecule in affinity assays, esp. for detecting large macromolecules such as ferritin.

PS Claim 51; Page 27; 156pp; English.

XX This sequence represents a synthetic ferritin peptide which was used as a functional surrogate in the conjugate of the invention. The novel CC labelled conjugate comprises at least one label attached to a functional CC surrogate of an analyte of interest. The surrogate is capable of CC competing effectively with the analyte for a limiting amount of an CC affinity receptor for the analyte. The conjugate exhibits an activity CC that is altered upon interaction with the affinity receptor and this CC activity can be measured and related to the amount of the analyte present in a sample. Functional surrogates such as this have an immunoreactive CC group that allows the surrogate to compete effectively and with the CC analyte for a limiting amount of its affinity receptor. Functional CC surrogates are able to mimic naturally occurring analytes. They can CC be labelled for use in standard competitive affinity assays (esp. CC homogeneous immunoassays) for detecting large macromolecules such as CC polypeptides, polysaccharides, polynucleotides, glycoproteins and CC lipid-containing macromolecules, as well as small haptens. Typical CC diagnostic analytes for detection include cardiac or tumour markers, CC allergens, hormones related to fertility-pregnancy or analytes CC associated with infectious disease. In particular, the assays are CC useful for detecting ferritin, follicle stimulating hormone, human CC growth hormone, immunoglobulin E, prolactin, parathyroid hormone, human CC placental lactogen, hepatitis antigens or antibodies against them, human CC chorionic gonadotropin, human luteinising hormone, cytomegalovirus, CC Chlamydia, Streptococcus A, rubella, toxoplasma, herpes virus, DK-MB, CC myoglobin, myosin light chain, troponin, carcinoembryonic antigen, CC alpha-fetoprotein, prostate-specific antigen and CA125 (a tumour marker). CC (Updated on 25-MAR-2003 to correct PI field.)

SQ Sequence 6 AA;

Query Match 94.7%; Score 18; DB 18; Length 6;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSLF 5
|||

DB 2 SSLF 5

RESULT 10
AAW10778
ID AAW10778 standard; peptide; 7 AA.

XX AAW10778;
AC AAW10778;
XX
DT 25-MAR-2003 (updated)
DT 10-NOV-1997 (first entry)

XX
DE Ferritin motif #21 important for selective binding affinity.

XX Functional surrogate; analyte; affinity receptor; immunoreactive group;
KW mimic; homogeneous immunoassay; detection; diagnostic analyte; Chlamydia;
KW cardiac marker; tumour marker; allergen; hormone; fertility; myoglobin;
KW pregnancy; infectious disease; ferritin; myosin light chain; troponin;
KW follicle stimulating hormone; human; growth hormone; immunoglobulin E;
KW prolactin; parathyroid hormone; placental lactogen; hepatitis antigen;
KW antibody; chorionic gonadotropin; luteinising hormone; cytomegalovirus;
KW Streptococcus; rubella; toxoplasma; DK-MB; prostate-specific antigen;
KW carcinoembryonic antigen; alpha-fetoprotein; herpes virus; CA125.

XX
OS Synthetic.
XX
XX
PN W09641172-A1.
XX
XX 19-DEC-1996.
PD
XX
PF 07-JUN-1996; 96WO-US10498.

XX
PR 07-JUN-1995; 95US-0476375.
XX
XX (CYTO-) CYTOGEN CORP.
PA
XX
XX Lee-owen FV, Carter JM;
PI
DR WPI; 1997-077284/07.
XX
XX
XX
PS Claim 56; Page 56; 156pp; English.

XX This sequence represents a peptide motif derived from ferritin which is CC important for selective binding affinity. Peptides containing motifs CC such as this may be used as functional surrogates in the conjugate of CC the invention. The novel labelled conjugate comprises at least one CC label attached to a functional surrogate of an analyte of interest. CC The surrogate is capable of competing effectively with the analyte CC for a limiting amount of an affinity receptor for the analyte. The CC conjugate exhibits an activity that is altered upon interaction with CC the affinity receptor and this activity can be measured and related to CC the amount of the analyte present in a sample. Functional surrogates CC such as this have an immunoreactive group that allows the surrogate to CC compete effectively and with the analyte for a limiting amount of its CC affinity receptor. Functional surrogates are able to mimic naturally CC occurring analytes. They can be labelled for use in standard competitive CC affinity assays (esp. homogeneous immunoassays) for detecting large CC macromolecules such as polypeptides, polysaccharides, polynucleotides, CC glycoproteins and lipid-containing macromolecules, as well as small CC haptens. Typical diagnostic analytes for detection include cardiac or CC tumour markers, allergens, hormones related to fertility-pregnancy or CC analytes associated with infectious disease. In particular, the assays CC are useful for detecting ferritin, follicle stimulating hormone, human CC growth hormone, immunoglobulin E, prolactin, parathyroid hormone, human CC placental lactogen, hepatitis antigens or antibodies against them, human CC chorionic gonadotropin, human luteinising hormone, cytomegalovirus, CC Chlamydia, Streptococcus A, rubella, toxoplasma, herpes virus, DK-MB, CC myoglobin, myosin light chain, troponin, carcinoembryonic antigen, CC alpha-fetoprotein, prostate-specific antigen and CA125 (a tumour marker). CC (Updated on 25-MAR-2003 to correct PI field.)

SQ Sequence 7 AA;

Query Match 94.7%; Score 18; DB 18; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSLF 5
|||

DB 3 SSLF 6

RESULT 11
AAW10779
ID AAW10779 standard; peptide; 7 AA.

XX AAW10779;
AC AAW10779;
XX
DT 25-MAR-2003 (updated)
DT 10-NOV-1997 (first entry)

XX
DE Ferritin motif #22 important for selective binding affinity.

XX Functional surrogate; analyte; affinity receptor; immunoreactive group;
KW mimic; homogeneous immunoassay; detection; diagnostic analyte; Chlamydia;
KW cardiac marker; tumour marker; allergen; hormone; fertility; myoglobin;
KW pregnancy; infectious disease; ferritin; myosin light chain; troponin;
KW follicle stimulating hormone; human; growth hormone; immunoglobulin E;
KW prolactin; parathyroid hormone; placental lactogen; hepatitis antigen;
KW antibody; chorionic gonadotropin; luteinising hormone; cytomegalovirus;

KM Streptococcus; rubella; toxoplasma; DK-MB; prostate-specific antigen;
 XX carcinoma; embryonic antigen; alpha-fetoprotein; herpes virus; CA125.
 OS Synthetic.
 XX WO9641172-A1.
 XX 19-DEC-1996.
 XX 07-JUN-1996; 96WO-US10498.
 XX 07-JUN-1995; 95US-0476375.
 XX (CYTO-) CYTOGEN CORP.
 XX Lee-owen FV, Carter JM;
 XX WPI; 1997-077284/07.
 XX Labelled functional surrogate of an analyte - useful as competitor
 PT molecule in affinity assays, esp. for detecting large macromolecules
 PT such as ferritin
 XX
 PS Claim 56; Page 56; 156pp; English.
 XX This sequence represents a peptide motif derived from ferritin which is
 CC important for selective binding affinity. Peptides containing motifs
 CC such as this may be used as functional surrogates in the conjugate of
 CC the invention. The novel labelled conjugate comprises at least one
 CC label attached to a functional surrogate of an analyte of interest.
 CC The surrogate is capable of competing effectively with the analyte
 CC for a limiting amount of an affinity receptor for the analyte. The
 CC conjugate exhibits an activity that is altered upon interaction with
 CC the affinity receptor and this activity can be measured and related to
 CC the amount of the analyte present in a sample. Functional surrogates
 CC such as this have an immunoreactive group that allows the surrogate to
 CC compete effectively and with the analyte for a limiting amount of its
 CC affinity receptor. Functional surrogates are used to mimic naturally
 CC occurring analytes. They can be labelled for use in standard competitive
 CC affinity assays (esp. homogeneous immunoassays) for detecting large
 CC macromolecules such as polypeptides, polysaccharides, polynucleotides,
 CC glycoproteins and lipid-containing macromolecules, as well as small
 CC haptens. Typical diagnostic analytes for detection include cardiac or
 CC tumour markers, allergens, hormones related to fertility-pregnancy or
 CC analyses associated with infectious disease. In particular, the assays
 CC are useful for detecting ferritin, follicle stimulating hormone, human
 CC growth hormone, immunoglobulin E, prolactin, parathyroid hormone, human
 CC placental lactogen, hepatitis antigens or antibodies against them, human
 CC chorionic gonadotropin, human luteinising hormone, cytomegalovirus,
 CC Chlamydia, Streptococcus a, rubella, toxoplasma, herpes virus, DK-MB,
 CC myoglobin, myosin light chain, tropoin, carcinoma; embryonic antigen,
 CC alpha-fetoprotein, prostate-specific antigen and CA125 (a tumour marker).
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 7 AA;
 XX
 QY Query Match 94.7%; Score 18; DB 18; Length 7;
 XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 DB Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX Ferritin motif #27 important for selective binding affinity.
 DE Functional surrogate; analyte; affinity receptor; immunoreactive group;
 XX mimic; homogeneous immunoassay; detection; diagnostic analyte; Chlamydia;
 XX cardiac marker; tumour marker; allergen; hormone; fertility; myoglobin;
 XX pregnancy; infectious disease; ferritin; myosin light chain; tropoin;
 XX follicle stimulating hormone; human; growth hormone; immunoglobulin E;
 XX prolactin; parathyroid hormone; placental lactogen; hepatitis antigen;
 XX antibody; chorionic gonadotropin; luteinising hormone; cytomegalovirus;
 XX Streptococcus; rubella; toxoplasma; DK-MB; prostate-specific antigen;
 XX carcinoma; embryonic antigen; alpha-fetoprotein; herpes virus; CA125.
 XX Synthetic.
 XX WO9641172-A1.
 XX 19-DEC-1996.
 XX 07-JUN-1996; 96WO-US10498.
 XX 07-JUN-1995; 95US-0476375.
 XX (CYTO-) CYTOGEN CORP.
 XX Lee-owen FV, Carter JM;
 XX WPI; 1997-077284/07.
 XX Labelled functional surrogate of an analyte - useful as competitor
 PT molecule in affinity assays, esp. for detecting large macromolecules
 PT such as ferritin
 XX
 PS Claim 51; Page 56; 156pp; English.
 XX This sequence represents a peptide motif derived from ferritin which is
 CC important for selective binding affinity. Peptides containing motifs
 CC such as this may be used as functional surrogates in the conjugate of
 CC the invention. The novel labelled conjugate comprises at least one
 CC label attached to a functional surrogate of an analyte of interest.
 CC The surrogate is capable of competing effectively with the analyte
 CC for a limiting amount of an affinity receptor for the analyte. The
 CC conjugate exhibits an activity that is altered upon interaction with
 CC the affinity receptor and this activity can be measured and related to
 CC the amount of the analyte present in a sample. Functional surrogates
 CC such as this have an immunoreactive group that allows the surrogate to
 CC compete effectively and with the analyte for a limiting amount of its
 CC affinity receptor. Functional surrogates are able to mimic naturally
 CC occurring analytes. They can be labelled for use in standard competitive
 CC affinity assays (esp. homogeneous immunoassays) for detecting large
 CC macromolecules such as polypeptides, polysaccharides, polynucleotides,
 CC glycoproteins and lipid-containing macromolecules, as well as small
 CC haptens. Typical diagnostic analytes for detection include cardiac or
 CC tumour markers, allergens, hormones related to fertility-pregnancy or
 CC analyses associated with infectious disease. In particular, the assays
 CC are useful for detecting ferritin, follicle stimulating hormone, human
 CC growth hormone, immunoglobulin E, prolactin, parathyroid hormone, human
 CC placental lactogen, hepatitis antigens or antibodies against them, human
 CC chorionic gonadotropin, human luteinising hormone, cytomegalovirus,
 CC Chlamydia, Streptococcus a, rubella, toxoplasma, herpes virus, DK-MB,
 CC myoglobin, myosin light chain, tropoin, carcinoma; embryonic antigen,
 CC alpha-fetoprotein, prostate-specific antigen and CA125 (a tumour marker).
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 7 AA;
 XX
 QY Query Match 94.7%; Score 18; DB 18; Length 7;
 XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 DB Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 13
AAW10801
ID AAW10801 standard; peptide, 7 AA.
XX
AC AAW10801;
XX
DT 25-MAR-2003 (updated)
DT 10-NOV-1997 (first entry)
XX
DE Ferritin motif #44 important for selective binding affinity.
XX
KW Functional surrogate; analyte; affinity receptor; immunoreactive group;
KW mimic; homogenous immunoassay; detection; diagnostic analyte; Chlamydia;
KW cardiac marker; tumour marker; allergen; hormone; fertility; myoglobin;
KW pregnancy; infectious disease; ferritin; myosin light chain; tropoin;
KW follicle stimulating hormone; human; growth hormone; immunoglobulin E;
KW prolactin; parathyroid hormone; placental lactogen; hepatitis antigen;
KW antibody; chorionic gonadotropin; luteinising hormone; cytomegalovirus;
KW streptococcus; rubella; toxoplasma; DK-MB; prostate-specific antigen;
KW carcinoembryonic antigen; alpha-fetoprotein; herpes virus; CA125.
XX
OS Synthetic.
XX
FN WO9641172-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US0498.
XX
PR 07-JUN-1995; 95US-0476375.
XX
PA (CYTO-) CYTOGEN CORP.
XX
PI Lee-owen FV, Carter JM;
XX
DR WPI; 1997-077284/07.
XX
PT Labelled functional surrogate of an analyte - useful as competitor
PT molecule in affinity assays, esp. for detecting large macromolecules
XX
PS Claim 56; Page 56; 156pp; English.
XX
CC This sequence represents a peptide motif derived from ferritin which is
CC important for selective binding affinity. Peptides containing motifs
CC such as this may be used as functional surrogates in the conjugate of
CC the invention. The novel labelled conjugate comprises at least one
CC label attached to a functional surrogate of an analyte of interest.
CC The surrogate is capable of competing effectively with the analyte
CC for a limiting amount of an affinity receptor for the analyte. The
CC conjugate exhibits an activity that is altered upon interaction with
CC the affinity receptor and this activity can be measured and related to
CC the amount of the analyte present in a sample. Functional surrogates
CC such as this have an immunoreactive group that allows the surrogate to
CC compete effectively and with the analyte for a limiting amount of its
CC affinity receptor. Functional surrogates are able to mimic naturally
CC occurring analytes. They can be labelled for use in standard competitive
CC affinity assays (esp. homogenous immunoassays) for detecting large
CC macromolecules such as polypeptides, polysaccharides, polynucleotides,
CC glycoproteins and lipid-containing macromolecules, as well as small
CC haptens. Typical diagnostic analytes for detection include cardiac or
CC tumour markers, allergens, hormones related to fertility-pregnancy or
CC analytes associated with infectious disease. In particular, the assays
CC are useful for detecting ferritin, follicle stimulating hormone, human
CC growth hormone, immunoglobulin E, prolactin, parathyroid hormone, human
CC placental lactogen, hepatitis antigens or antibodies against them, human
CC chorionic gonadotropin, human luteinising hormone, cytomegalovirus,
CC Chlamydia, streptococcus a, rubella, toxoplasma, herpes virus, DK-MB,
CC myoglobin, myosin light chain, tropoin, carcinoembryonic antigen,
CC alpha-fetoprotein, prostate-specific antigen and CA125 (a tumour marker).
CC (Updated on 25-MAR-2003 to correct PI field.)

XX
SQ Sequence 7 AA;
XX
Query Match 94.7%; Score 18; DB 18; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2 SSF 5
Db 3 SSF 6
XX
RESULT 14
AAW10802
ID AAW10802 standard; peptide, 7 AA.
XX
AC AAW10802;
XX
DT 25-MAR-2003 (updated)
DT 10-NOV-1997 (first entry)
XX
DE Ferritin motif #45 important for selective binding affinity.
XX
KW Functional surrogate; analyte; affinity receptor; immunoreactive group;
KW mimic; homogenous immunoassay; detection; diagnostic analyte; Chlamydia;
KW cardiac marker; tumour marker; allergen; hormone; fertility; myoglobin;
KW pregnancy; infectious disease; ferritin; myosin light chain; tropoin;
KW follicle stimulating hormone; human; growth hormone; immunoglobulin E;
KW prolactin; parathyroid hormone; placental lactogen; hepatitis antigen;
KW antibody; chorionic gonadotropin; luteinising hormone; cytomegalovirus;
KW streptococcus; rubella; toxoplasma; DK-MB; prostate-specific antigen;
KW carcinoembryonic antigen; alpha-fetoprotein; herpes virus; CA125.
XX
OS Synthetic.
XX
FN WO9641172-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US0498.
XX
PR 07-JUN-1995; 95US-0476375.
XX
PA (CYTO-) CYTOGEN CORP.
XX
PI Lee-owen FV, Carter JM;
XX
DR WPI; 1997-077284/07.
XX
PT Labelled functional surrogate of an analyte - useful as competitor
PT molecule in affinity assays, esp. for detecting large macromolecules
XX
PS Claim 56; Page 56; 156pp; English.
XX
CC This sequence represents a peptide motif derived from ferritin which is
CC important for selective binding affinity. Peptides containing motifs
CC such as this may be used as functional surrogates in the conjugate of
CC the invention. The novel labelled conjugate comprises at least one
CC label attached to a functional surrogate of an analyte of interest.
CC The surrogate is capable of competing effectively with the analyte
CC for a limiting amount of an affinity receptor for the analyte. The
CC conjugate exhibits an activity that is altered upon interaction with
CC the affinity receptor and this activity can be measured and related to
CC the amount of the analyte present in a sample. Functional surrogates
CC such as this have an immunoreactive group that allows the surrogate to
CC compete effectively and with the analyte for a limiting amount of its
CC affinity receptor. Functional surrogates are able to mimic naturally
CC occurring analytes. They can be labelled for use in standard competitive
CC affinity assays (esp. homogenous immunoassays) for detecting large
CC macromolecules such as polypeptides, polysaccharides, polynucleotides,
CC glycoproteins and lipid-containing macromolecules, as well as small
CC haptens. Typical diagnostic analytes for detection include cardiac or

CC tumour markers, allergens, hormones related to fertility-pregnancy or
 CC analyses associated with infectious disease. In particular, the assays
 CC are useful for detecting ferritin, follicle stimulating hormone, human
 CC growth hormone, immunoglobulin E, prolactin, parathyroid hormone, human
 CC placental lactogen, hepatitis antigens or antibodies against them, human
 CC chorionic gonadotropin, human luteinizing hormone, cytomegalovirus,
 CC Chlamydia, Streptococcus A, Rubella, Toxoplasma, Herpes virus, DK-WB,
 CC myoglobin, myosin light chain, tropoin, carcinoembryonic antigen,
 CC alpha-fetoprotein, prostate-specific antigen and CA125 (a tumour marker).
 CC (Updated on 25-MAR-2003 to correct PI field.)
 CC
 XX
 SQ Sequence 7 AA;
 Query Match 94.7%; Score 18; DB 18; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 SSLF 5
 Db 3 SSLF 6
 RESULT 15
 ABG96785
 ID ABG96785 standard; Peptide; 7 AA.
 XX
 AC ABG96785;
 XX
 DT 16-DEC-2002 (first entry)
 XX
 DE Human leukocyte antigen (HLA) B*1501 epitope #59.
 XX
 KW Soluble human leukocyte antigen; HLA; SHLA; cell pharm;
 KW multimeric HLA complex; cytotoxic T lymphocytes; CTL; immune response;
 KW viral infection; cancer; autoimmune disease; vaccine development; MHC;
 KW major histocompatibility complex; diagnostic development;
 KW HLA class I polymorphism; HLA-B15 allele.
 XX
 OS Homo sapiens.
 XX
 PN W0200262846-A2.
 XX
 PD 15-AUG-2002.
 XX
 PF 18-DEC-2001; 2001MO-US49744.
 XX
 PR 18-DEC-2000; 2000US-256409P.
 PR 18-DEC-2000; 2000US-256410P.
 PR 10-OCT-2001; 2001US-0974366.
 XX
 PA (HILD/) HILDEBRAND W H.
 PA (PRIL/) PRILLIMAN K R.
 XX
 PI Hildebrand WH, Prilliman KR;
 DR WPI, 2002-698563/75.
 XX
 PT Producing soluble human leukocyte antigen (HLA) in cell pharm useful
 PT for studies of peptide loading for characterizing human immune
 PT responses involves using HLA allelic CDNA or genomic DNA as starting
 PT material
 XX
 PS Disclosure, Page 172; 300pp; English.
 XX
 CC The invention describes a method of producing soluble human leukocyte
 CC antigen (HLA) molecules (SHLA) in cell pharm involving amplifying HLA
 CC allelic DNA by PCR using a locus specific primer to produce truncated a
 CC PCR product (PI), inserting PI into mammalian expression vector;
 CC electroporating the plasmid into a host cell; inoculating the cell pharm
 CC with the host cell such that cell pharm produces SHLA. A multimeric HLA
 CC complex (I) is useful for testing functionality of peptide ligands bound
 CC by at least two soluble HLA molecules. (I) can be tested for its ability
 CC to serve as ligands for cytotoxic T lymphocytes (CTLs) and induce immune

CC responses in humans. (I) is useful for studying T cell responses to
 CC pathological conditions such as viral infections and cancer, and for
 CC modulating the human immune system to induce tolerance in autoimmune
 CC diseases. The individual secreted major histocompatibility complex (MHC)
 CC molecules produced are useful for studies of peptide loading (i.e., in
 CC vaccine development) and to the development of diagnostics. With the
 CC secreted MHC molecules, naturally loaded peptides can be eluted from the
 CC MHC molecule and characterized. The secreted MHC molecules allow the
 CC assessment of structural and functional impact of HLA class I
 CC polymorphism. The molecules are also useful to generate ligands and hence
 CC ligand maps from the peptide pools extracted from series of distinct yet
 CC related class I HLA-B15 alleles; compare the different ligand maps to
 CC identify potentially shared elements; and characterize the elements
 CC identified to positively or negatively validate the occurrence of
 CC overlapping ligands. The truncated version of (MHC) can be produced in
 CC mammalian or insect/bacterial cells such that milligram or greater
 CC quantities of an individual class I or class II molecule can be obtained.
 CC This sequence represents a HLA (human leukocyte antigen) epitope.
 CC
 XX
 SQ Sequence 7 AA;
 Query Match 94.7%; Score 18; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 SSLF 5
 Db 4 SSLF 7

Search completed: January 29, 2004, 14:52:14
 Job time : 20.4853 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:46:49 ; Search time 6.17647 Seconds

(without alignments)
77.851 Million cell updates/sec

Title: US-10-032-950-8

Perfect score: 19

Sequence: 1 XSLF 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 76:*

2: PIR1:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	94.7	18	2	S57518
2	18	94.7	18	2	A60739
3	18	94.7	21	2	S03481
4	18	94.7	24	2	PH1696
5	18	94.7	24	2	PH1697
6	18	94.7	29	2	E47719
7	18	94.7	34	2	T11814
8	18	94.7	46	2	B90821
9	18	94.7	47	2	C89995
10	18	94.7	54	2	S34728
11	18	94.7	54	2	T06514
12	18	94.7	55	2	B71683
13	18	94.7	59	2	S01877
14	18	94.7	62	2	T07470
15	18	94.7	62	2	H21168
16	18	94.7	65	2	P00719
17	18	94.7	66	2	C82699
18	18	94.7	70	2	A84033
19	18	94.7	71	2	T18055
20	18	94.7	71	2	E82025
21	18	94.7	73	2	T25649
22	18	94.7	74	2	E82864
23	18	94.7	75	2	A81539
24	18	94.7	76	2	T17719
25	18	94.7	77	2	A86399
26	18	94.7	78	2	T03930
27	18	94.7	79	2	T02948
28	18	94.7	80	2	D33285
29	18	94.7	80	2	C81003

ALIGNMENTS

30	18	94.7	80	2	S19984	hypothetical prote
31	18	94.7	80	2	S19987	hypothetical prote
32	18	94.7	82	2	C69013	hypothetical prote
33	18	94.7	83	2	C84274	hypothetical prote
34	18	94.7	84	2	AB2604	hypothetical prote
35	18	94.7	85	2	C82369	this protein VC006
36	18	94.7	92	1	QOECXP	hypothetical prote
37	18	94.7	93	2	G64521	hypothetical prote
38	18	94.7	93	2	C71984	hypothetical prote
39	18	94.7	95	2	AG0277	probable phage-rel
40	18	94.7	97	2	A83971	hypothetical prote
41	18	94.7	99	2	P88105	conserved hypothet
42	18	94.7	100	2	S44892	ZK112.4 protein -
43	18	94.7	101	2	S73700	probable lipoprote
44	18	94.7	101	2	T38892	very hypothetical
45	18	94.7	102	2	G71025	hypothetical prote

RESULT 1

S57518
T cell receptor beta chain V region - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 10-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1995
C/Accession: S57518
R/Butrows, S.R.; Sallins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Aagaet, V.P.
submitted to the EMBL Data Library, June 1995
A/Description: T cell receptor repertoire for a viral epitope in humans is diversified
A/Reference number: S57494
A/Accession: S57518
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-18 <BUR>
A/Cross-references: EMBL:249920; NID:G887490; PIDN:CAA90166.1; PID:G887491
C/Keywords: T-cell receptor

Query Match

Best Local Similarity 94.7%; Score 18; DB 2; Length 18;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
DB 3 SSLF 6

RESULT 2

A60739
somatotropin - guinea pig (fragment)
N/Alternate names: growth hormone
C/Species: Cavia porcellus (guinea pig)
C/Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 17-Mar-1999
C/Accession: A60739
R/Gabrielsson, B.; Fairhall, K.M.; Robinson, I.C.A.F.
J. Endocrinol. 124, 371-380, 1990
A/Title: Growth hormone secretion in the guinea pig.
A/Reference number: A60739; MUID:90237710; EMBL:1970601
A/Accession: A60739
A/Molecule type: protein
A/Residues: 1-20 <GAB>
C/Comment: The growth hormone system in the guinea pig is unusual among mammals in th
C/Superfamily: prolactin
C/Keywords: anterior pituitary; growth factor; hormone

Query Match 94.7%; Score 18; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
DB 7 SSLF 10

RESULT 3

T-cell receptor alpha chain J region (TA20) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 05-Nov-1999
 C:Accession: S03481
 R:Ardan, B.; Klotz, J.L.; Siu, G.; Hood, L.E.
 Nature 316, 783-787, 1995
 A:Title: Diversity and structure of genes of the alpha family of mouse T-cell antigen re
 A:Reference number: S03467; MUID:85286324; PMID:3839904
 A:Accession: S03461
 A:Molecule type: mRNA
 A:Residues: 1-21 <ARD>
 A:Cross-references: EMBL:X02975; NID:g54392; PIDN:CA26717.1; PID:9773240
 A:Note: this sequence was determined from the differentiated gene
 C:Keywords: T-cell receptor
 F:1-17/Domain: T-cell receptor alpha chain J region #status predicted <JTB>

Query Match

Best Local Similarity 94.7%; Score 18; DB 2; Length 21;
 Pred. No. 75;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5

Db 2 SSUF 5

RESULT 4

Ig heavy chain V region (clone NP-7-9) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
 C:Accession: PH1696
 R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
 J. Exp. Med. 178, 295-307, 1993
 A:Title: Antigen-driven B cell differentiation in vivo.
 A:Reference number: PH1675; MUID:93301607; PMID:8315385
 A:Accession: PH1666
 A:Molecule type: mRNA
 A:Residues: 1-24 <MCH>
 A:Experimental source: B cell
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotrimer; immunoglobulin

Query Match

Best Local Similarity 94.7%; Score 18; DB 2; Length 24;
 Pred. No. 86;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5

Db 19 SSUF 22

RESULT 5

Ig heavy chain V region - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
 C:Accession: PH1697; PH1699
 R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
 J. Exp. Med. 178, 295-307, 1993
 A:Title: Antigen-driven B cell differentiation in vivo.
 A:Reference number: PH1675; MUID:93301607; PMID:8315385
 A:Accession: PH1697
 A:Molecule type: mRNA
 A:Residues: 1-24 <MCH>
 A:Experimental source: B cell; clone NP-7-10
 A:Note: the authors translated the codon GCT for residue 23 as Asp
 A:Accession: PH1699
 A:Molecule type: mRNA
 A:Residues: 1-10, 'A', '12-13, 'R', '15-20, 'T', '22, 'D', '24 <MC2>
 A:Experimental source: B cell; clone NP-7-12

C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotrimer; immunoglobulin

Query Match 94.7%; Score 18; DB 2; Length 24;
 Best Local Similarity 100.0%; Pred. No. 86;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5

Db 19 SSUF 22

RESULT 6

house-dust-mite-reactive T-cell receptor beta chain (CD4+ clone DE26, V(D)J junctiona
 C:Species: Homo sapiens (man)
 C:Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997
 C:Accession: B47719
 R:Wiederdorn, L.R.; O'Heir, R.E.; Hewitt, C.R.; Lamb, J.R.; Owen, M.J.
 Proc. Natl. Acad. Sci. U.S.A. 90, 8214-8218, 1993
 A:Title: In vivo clonal dominance and limited T-cell receptor usage in human CD4+ T-c
 A:Reference number: A47719; MUID:93376774; PMID:8367485
 A:Accession: B47719
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-29 <WBD>
 A:Note: sequence extracted from NCBI backbone (NCBIN:137825, NCBI:137831)
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: T-cell receptor

Query Match

Best Local Similarity 94.7%; Score 18; DB 2; Length 29;
 Pred. No. 1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5

Db 3 SSUF 6

RESULT 7

hypothetical protein ORF34a - Norway spruce chloroplast
 C:Species: chloroplast Picea abies (Norway spruce)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
 C:Accession: T11814
 R:Kluemper, S.; Kanka, S.; Riesner, D.; Ertseid, M.
 submitted to the EMBL Data Library, March 1997
 A:Description: Characterisation of a Norway spruce chloroplast DNA clone: Complete nu
 A:Reference number: Z17349
 A:Accession: T11814
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-34 <KLU>
 A:Cross-references: EMBL:U92462; NID:g2959581; PID:g2959590
 A:Genetics:
 C:Keywords: chloroplast

Query Match

Best Local Similarity 94.7%; Score 18; DB 2; Length 34;
 Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5

Db 13 SSUF 16

RESULT 8

hypothetical protein Ecs158 (imported) - Escherichia coli (strain C157.H7, substrain
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C:Accession: B90821

R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A/Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene
 A/Reference number: A99623; NCID:21156231; PMID:11258796
 A/Accession: B90821
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-46 <HAY>
 A/Cross-references: GB:BA000007; PIDN:BA034961.1; PID:G13361002; GSPDB:GN00154
 A/Experimental source: strain O157:H7, substrain RIMD 0509952
 C/Genetics:
 A/Gene: ECs1538

Query Match 94.7%; Score 18; DB 2; Length 46;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
 ||||
 Db 41 SSLF 44

RESULT 9
 C89995
 AgRP protein [imported] - *Staphylococcus aureus* (strain N315)
 C/Species: *Staphylococcus aureus*
 C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 A/Accession: C89995
 R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud
 ma, A.; Mizumachi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 351, 1225-1240, 2001
 A/Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A/Reference number: A89758; NCID:2111952; PMID:11418146
 A/Accession: C89995
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-47 <KUR>
 A/Cross-references: GB:BA000018; PID:G13701831; PIDN:BA034324.1; GSPDB:GN00149
 A/Experimental source: strain N315
 C/Genetics:
 A/Gene: agrD

Query Match 94.7%; Score 18; DB 2; Length 47;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
 ||||
 Db 29 SSLF 32

RESULT 10
 S34728
 ribose-phosphate diphosphokinase (EC 2.7.6.1) - *Listeria ivanovii* (fragment)
 C/Species: *Listeria ivanovii*
 C/Date: 02-Dec-1993 #sequence_revision 25-Oct-1996 #text_change 03-Jun-2002
 A/Accession: S54405; S34728
 R/Lampidis, R.; Gross, R.; Socolovic, Z.; Goebel, W.; Krefl, J.
 Mol. Microbiol. 13, 141-151, 1994
 A/Title: The virulence regulator protein of *Listeria ivanovii* is highly homologous to Pr
 A/Reference number: S54403; NCID:95075291; PMID:7984088
 A/Accession: S54405
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-54 <LAM>
 A/Cross-references: EMBL:X72685; NID:G396074; PIDN:CA051232.1; PID:G396077
 A/Experimental source: ATCC 19119
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1993
 C/Genetics:
 A/Gene: prs
 C/Function:

A/Pathway: nucleotide biosynthesis
 C/Superfamily: ribose-phosphate pyrophosphokinase catalytic chain
 C/Keywords: diphosphotransferase; magnesium; nucleotide biosynthesis

Query Match 94.7%; Score 18; DB 2; Length 54;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
 ||||
 Db 50 SSLF 53

RESULT 11
 T06514
 hypothetical protein - garden pea
 C/Species: *Pisum sativum* (garden pea)
 C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 08-Oct-1999
 A/Accession: T06514
 R/Carrasco, F.; Marco, F.
 submitted to the EMBL Data Library, June 1996
 A/Description: pea S-adenosylmethionine decarboxylase.
 A/Reference number: 215729
 A/Accession: T06514
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: mRNA
 A/Residues: 1-54 <CAR>
 A/Cross-references: EMBL:U060592; NID:G1421750; PIDN:AA03864.1; PID:G1421751
 A/Experimental source: cultivar Alaska

Query Match 94.7%; Score 18; DB 2; Length 54;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
 ||||
 Db 17 SSLF 20

RESULT 12
 B71683
 hypothetical protein RP279 - *Rickettsia prowazekii*
 C/Species: *Rickettsia prowazekii*
 C/Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
 A/Accession: B71683
 R/Andersson, S.G.B.; Zomrodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark
 Nature 396, 133-140, 1998
 A/Title: The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria.
 A/Reference number: A71630; NCID:99039499; PMID:9823893
 A/Accession: B71683
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-55 <AND>
 A/Cross-references: GB:AJ235271; GB:AJ235269; NID:G3868717; PIDN:CA14740.1; PID:G386
 A/Experimental source: strain Madrid E
 C/Genetics:
 A/Gene: RP279

Query Match 94.7%; Score 18; DB 2; Length 55;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
 ||||
 Db 14 SSLF 17

RESULT 13
 S01877
 NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - brine shrimp mitochondrion (f
 C/Species: mitochondrion *Artemia* sp. (brine shrimp)
 C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 03-Jun-2002
 A/Accession: S01877

R:Batuecas, B.; Garesse, R.; Calleja, M.; Valverde, J.R.; Marco, R.
 Nucleic Acids Res. 15, 6515-6529, 1988
 A>Title: Genome organization of *Artemia* mitochondrial DNA.
 A:Reference number: S01207; MUID:88289417; PMID:3135541
 A:Accession: S01877
 A:Molecule type: DNA
 A:Residues: 1-59 <BAT>
 A:Cross-references: EMBL:X07663; NID:G600442; PID:G600443
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC4
 C:Keywords: mitochondrion; NAD; oxidoreductase

Query Match 94.7%; Score 18; DB 2; Length 59;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 SSLF 5
 Db 30 SSLF 33

RESULT 14
 T07470
 hypocholesteric protein 62b - Japanese black pine chloroplast
 C:Species: chloroplast Pinus thunbergiana (Japanese black pine)
 C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 18-Aug-2000
 C:Accession: T07470
 R:Wakasugi, T.; Teudzuki, J.; Ito, S.; Nakashima, K.; Teudzuki, T.; Sugitara, M.
 Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994
 A>Title: Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome
 A:Reference number: 216030; MUID:95024047; PMID:7937893
 A:Accession: T07470
 A>Status: preliminary; translated from GB/EMBL/DDBB
 A:Molecule type: DNA
 A:Residues: 1-62 <WAK>
 A:Cross-references: EMBL:D17510; NID:G529643; PID:BA04348.1; PID:G1262631
 C:Genetics:
 A:Genome: chloroplast
 C:Keywords: chloroplast

Query Match 94.7%; Score 18; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 SSLF 5
 Db 56 SSLF 59

RESULT 15
 H72168
 A43R protein - variola minor virus (strain Garcia-1966)
 C:Species: variola minor virus
 C>Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999
 C:Accession: H72168
 R:Shchelkunov, S.N.; Tormenin, A.V.; Gutorov, V.V.; Safonov, P.F.; Massung, R.F.; Lopar
 submitted to GenBank, March 1998
 A>Description: Analysis of the complete coding sequence of DNA of *Alastrim variola minor*
 A:Reference number: A72150
 A:Accession: H72168
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-62 <SHC>
 A:Cross-references: GB:Y16780; NID:G5830555; PID:CA854746.1; PID:el542702; PID:G5830707
 A:Experimental source: strain Garcia-1966
 C:Genetics:
 A:Gene: A43R

Query Match 94.7%; Score 18; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
 Db 5 SSLF 8

Search completed: January 29, 2004, 14:58:25
 Job time: 7.17647 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:45:44 ; Search time 3.23529 Seconds

(without alignments)
72.678 Million cell updates/sec

Title: US-10-032-950-8

Perfect score: 19

Sequence: 1 XSLF 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	94.7	55	Y279_RICPR	Q94888 rickettsia
2	18	94.7	59	NUSM_ARTSA	P19047 attemia sal
3	18	94.7	62	PEPL_PINTH	P52805 pinus thunb
4	18	94.7	75	EX7S_CHLPP	Q94194 chlamydia p
5	18	94.7	77	SECG_CORG	Q94469 corynebacte
6	18	94.7	79	CYB_DIPCA	P16359 dipodomys c
7	18	94.7	80	YCU5_OENV	P31556 oenothera v
8	18	94.7	100	YOG4_CABEL	P34613 caenorhabd
9	18	94.7	101	YAYC_SCHPO	Q10219 schistosach
10	18	94.7	101	YB67_MYCPN	P75316 mycoplasma
11	18	94.7	103	V078_FOWPV	Q91542 fowlpox vir
12	18	94.7	104	YU22_YEAST	P47059 saccharomyc
13	18	94.7	105	R44P_CANMA	O00477 candida mal
14	18	94.7	105	RL12_SULAC	P08055 sulfolobus
15	18	94.7	112	PT17_STYPL	P28209 styela plic
16	18	94.7	114	YCR2_YEAST	P28620 saccharomyc
17	18	94.7	122	BE11_ARATH	Q94219 arabidopsis
18	18	94.7	124	HVID_HUMAN	P01760 homo sapien
19	18	94.7	124	UL39_HCMVA	P16813 human cytom
20	18	94.7	125	C560_PORPU	P80478 porphyra pu
21	18	94.7	125	YB3M_ECOLI	P75813 escherichia
22	18	94.7	127	C560_CROCK	P46934 chondrus cr
23	18	94.7	130	Y1B2_YEAST	P40551 saccharomyc
24	18	94.7	133	YB3U_YEAST	P38350 saccharomyc
25	18	94.7	135	TB11_MOUSE	P01734 mus musculu
26	18	94.7	140	RS19_SULSO	Q94343 sulfolobus
27	18	94.7	140	RS19_SULSO	Q94343 sulfolobus
28	18	94.7	141	YK13_YEAST	P16128 saccharomyc
29	18	94.7	145	RK18_PBA	P16591 pisum sativ
30	18	94.7	147	IL4_MESAU	Q60440 mesocricetu
31	18	94.7	148	YW25_YEAST	P40219 saccharomyc
32	18	94.7	149	RS19_METKA	O81WP2 methanopyru
33	18	94.7	154	SSRP_SYNY3	P74355 synechocyst

34	18	94.7	154	YWEA_BACSU	P39632 bacillus su
35	18	94.7	157	YXS2_PSEPU	Q05092 pseudomonas
36	18	94.7	157	YQAA_HAEIN	P44005 haemophilus
37	18	94.7	159	YML9_HELAN	P41248 helianthus
38	18	94.7	159	YML9_OENBE	P08768 oenothera b
39	18	94.7	161	YK19_CABEL	P34320 caenorhabd
40	18	94.7	162	YSAF_YERPE	Q56978 yerinia pe
41	18	94.7	164	GLB1_CHLEU	Q08753 chlamydomon
42	18	94.7	167	VP6_BPPI6	P11128 bacteriopho
43	18	94.7	170	YURS_YEAST	P46991 saccharomyc
44	18	94.7	172	YU6M_MOUSE	P03925 mus musculu
45	18	94.7	174	YU6M_DROME	P18993 drosophila

ALIGNMENTS

```

RESULT 1
Y279_RICPR          STANDARD;          PRT;          55 AA.
ID Y279_RICPR
AC Q9ZDP8;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein RP279.
GN RP279.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiase; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
RX MEDLINE=99039499; PubMed=9923893;
RA Andersson S.G.E.; Zomorodipour A.; Andersson J.O.;
RA Sichenhitz-Fonten T.; Alemark U.C.M.; Podowski R.M.; Naeslund A.K.;
RA Eriksson A.-S.; Winkler H.H.; Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria."
RL Nature 396:133-140(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC -----
DR EMBL; AJ235271; CAA14740.1; -.
DR PIR; B71683; B71683.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 55 AA; 6202 MW; CC6D7C51D29BD34E CRC64;
QY
Db 2 SLP 5
14 SLP 17
Query Match 94.7%; Score 18; DB 1; Length 55;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 2
ID NUSM_ARTSA          STANDARD;          PRT;          59 AA.
AC P19047;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3) (Fragment).
GN N5.
OS Artemia salina (brine shrimp).

```

```

OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
OC Artemiidae; Artemia.
OK NCBI_TaxID=85549;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88289417; PubMed=1135541;
RT Batuecas B., Garesse R., Calleja M., Valverde J.R., Marco R.;
RU "Genome organization of Artemia mitochondrial DNA.";
RU Nucleic Acids Res. 16:6515-6529(1988).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07663; CAA30510.1;
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
FT NON TER
SQ SEQUENCE 59 AA; 6585 MW; 60C9E5E04827DC23 CRC64;

Query Match          94.7%; Score 18; DB 1; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5
   |||
   30 SSUF 33

Db

RESULT 3
ID PETL_PINTH STANDARD; PRT; 62 AA.
AC P52805;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome b6-f complex subunit VI (Cytochrome b6f complex subunit
DE petl).
GN PETL.
OS Pinus thunbergii (Green pine) (Japanese black pine).
OC Chlorophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=3350;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95024047; PubMed=7937893;
RT Wakisugi T., Tsudzuki U., Ito S., Nakashima K., Tsudzuki T.,
RA Sugisue M.;
RT "Loss of all ndh genes as determined by sequencing the entire
RT chloroplast genome of the black pine Pinus thunbergii.";
RT Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).
CC -1- FUNCTION: THE CYTOCHROME B6-F COMPLEX FUNCTIONS IN THE LINEAR
CC CROSS-MEMBRANE TRANSPORT OF ELECTRONS BETWEEN PHOTOSYSTEM II AND
CC I, AS WELL AS IN CYCLIC ELECTRON FLOW AROUND PHOTOSYSTEM II. PETL
CC IS IMPORTANT FOR PHOTOTROPHIC GROWTH AS WELL AS FOR ELECTRON
CC TRANSFER EFFICIENCY AND STABILITY OF THE CYTOCHROME B6-F COMPLEX.
CC -1- SUBCELLULAR LOCATION: Thylakoid membrane-associated.
CC -1- SIMILARITY: Belongs to the petl family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; D17510; BAA04348.1;
DR PIR; T07470; T07470.
DR HAMAP; MF 00433; atypical; 1.
DR Pfam; PF05115; Petl; 1.
KW Electron transport; Chloroplast; Respiratory chain; Transmembrane;
KW Thylakoid.
FT TRANSMEM
SQ SEQUENCE 15 35 POTENTIAL.
   62 AA; 6924 MW; 85D87358A3D7E87F CRC64;

Query Match          94.7%; Score 18; DB 1; Length 62;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5
   |||
   56 SSUF 59

Db

RESULT 4
ID EX7S_CHLPH STANDARD; PRT; 75 AA.
AC Q9X1Y4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable exodeoxyribonuclease VII small subunit (EC 3.1.11.6)
DE (exonuclease VII small subunit).
GN XSEB OR CPN1061.1 OR CP0788.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RT Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uetters T., Berry K., Baas S.,
RA Lusher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., Deboy R., Kolony J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.W.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RT Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RT Nucleic Acids Res. 28:2311-2314(2000).
CC -1- FUNCTION: BIDIRECTIONALLY DEGRADES SINGLE-STRANDED DNA INTO LARGE
CC ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER
CC INTO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates. (By
CC similarity).
CC -1- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE XSEB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: AE001686; -; NOT_ANNOTATED_CDS.
 DR EMBL: AE002238; AAF38587.1; -;
 DR EMBL: AF002548; -; NOT_ANNOTATED_CDS.
 DR PIR: A81539; A81539.
 DR TIGR: CP0788; -;
 DR HAMAP: MF_00337; -; 1.
 DR InterPro: IPR003761; Exonuc_VII_S.
 DR Pfam: PF02609; Exonuc_VII_S; 1.
 DR ProDom: PD028235; Exonuc_VII_S; 1.
 DR TrEMBL: TIGR01280; xseB; 1.
 KM Hydroxylase; Nuclease; Exonuclease; Complete proteome.
 SQ SEQUENCE 75 AA; 8809 MW; AA05874AF8A3728 CRC64;

Query Match
 Best Local Similarity 94.7%; Score 18; DB 1; Length 75;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
 Db 65 SSLF 68

RESULT 5
 SEC0 CORGL STANDARD; PRT; 77 AA.
 AC 032459;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Protein-export membrane protein secG.
 GN SEC0 OR CG11584.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Jakoby M.J., Burkovski A., Meier-Magner J.;
 RT "Membrane uptake in Corynebacterium glutamicum is regulated on the
 RT level of expression and enzyme activity."
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
 RT Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT. PARTICIPATES IN AN EARLY
 CC -1- EVENT OF PROTEIN TRANSLOCATION (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE SEC0 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: AJ007732; CA07632.1; -;
 DR EMBL: AP005279; BAB98977.1; -;
 DR InterPro: IPR004692; SecG.
 DR Pfam: PF03840; SecG; 1.
 DR PRINTS: PR01651; SEC0EXPORT.
 DR TrEMBL: TIGR00810; secg; 1.
 KM Protein transport; Translocation; Transmembrane; Complete proteome.

FT TRANSMEM 2 22 POTENTIAL.
 FT TRANSMEM 55 75
 SQ SEQUENCE 77 AA; 8098 MW; B3819B9DACS0140C CRC64;

Query Match
 Best Local Similarity 94.7%; Score 18; DB 1; Length 77;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
 Db 31 SSLF 34

RESULT 6
 CYB_DIPCA STANDARD; PRT; 79 AA.
 AC P16359;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cytochrome b (Fragment).
 GN MT009 OR COB OR CYTB.
 OS Dipodomys californicus (Kangaroo rat).
 OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteleostomi; Rodentia; Sciurognathia; Heteromyidae;
 OC Dipodomys; Dipodomys.
 OX NCBI_TaxID=10017;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89345630; PubMed=2762322;
 RA Kocher T.D., Thomas W.K., Meyer A., Edwards S.V., Paabo S.,
 RA Villablanca F.X., Wilson A.C.;
 RT "Dynamics of mitochondrial DNA evolution in animals: amplification
 RT and sequencing with conserved primers."
 RL Proc. Natl. Acad. Sci. U.S.A. 86:6196-6200(1989).
 CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase
 CC complex (complex III or cytochrome b-c1 complex), which is a
 CC respiratory chain that generates an electrochemical potential
 CC coupled to ATP synthesis (By similarity).
 CC -1- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or
 CC b562) is low-potential and absorbs at about 562, and heme 2 (or BH
 CC or b566) is high-potential and absorbs at about 566 (By
 CC similarity).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 CC cytochrome c1 and the Rieske protein (By similarity).
 CC -1- SIMILARITY: Belongs to the cytochrome b family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: M25684; AAA31715.1; -;
 DR InterPro: IPR005798; Cytb_b6_C.
 DR InterPro: IPR005797; Cytb_b6_N.
 DR Pfam: PF00033; cytochrome_b_N; 1.
 DR PROSITE: PS00193; CYTOCHROME_B_Q0, PARTIAL.
 DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
 KM Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
 KM Heme.
 FT NON TER 1 1
 FT METAL 37 37 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 51 51 IRON 2 (HEME B566 AXIAL LIGAND).
 FT NON TER 79 79
 SQ SEQUENCE 79 AA; 9067 MW; 62F9B7E195C8E4EE CRC64;

Query Match
 Best Local Similarity 94.7%; Score 18; DB 1; Length 79;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSLF 5
 DB 41 SSLF 44

RESULT 7

YCL5_OENV1 STANDARD; PRT; 80 AA.

AC P31566; 01-JUL-1993 (Rel. 26, last sequence update)

DT 01-JUL-1993 (Rel. 26, last sequence update)

DE 16-OCT-2001 (Rel. 40, last annotation update)

GN Hypothetical 9.4 kDa protein YCL5 (ORF 80).

OS Oenothera villaricae, and Oenothera odorata).

OC Chloplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC Eucosids II; Myrtales; Onagraceae; Oenothera.

OX NCBI_TaxID=3941, 3946;

RN [1] SEQUENCE FROM N.A.

RP MEDLINE=93169690; PubMed=8435856;

RA Nimyck R., Shoendorf T., Hachrel W.;

RT "in-frame length mutations associated with short tandem repeats are

RT located in unassigned open reading frames of Oenothera chloroplast

RT CNA.";

RT Curr. Genet. 23:265-270(1993).

CC -1 SIMILARITY: BELONGS TO THE YCL5 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: X64615; CAA45897.2; -

DR EMBL: X64616; CAA45899.2; -

DR PIR: S19984; S19987.

DR PIR: S19987; S19987.

KW Chloroplast; Hypothetical protein.

SQ SEQUENCE 80 AA; 9408 MW; B5C0BCDD1D354846 CRC64;

Query Match 94.7%; Score 18; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSLF 5
 DB 6 SSLF 9

RESULT 8

YOG4_CAEEL STANDARD; PRT; 100 AA.

AC P34613; 01-FEB-1994 (Rel. 28, last sequence update)

DT 01-FEB-1994 (Rel. 28, last sequence update)

DE 28-FEB-2003 (Rel. 41, last annotation update)

GN Hypothetical protein ZK112.4 in chromosome III.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OC NCBI_TaxID=6239;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=Bristol N2;

RX MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Woldman P.;

RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

RT elegans.";

RL Nature 368:32-38(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: L14324; AAA28187.1; -

DR PIR: S44892; S44892.

DR Wormpep: ZK112.4; CE00375.

KW Hypothetical protein.

SQ SEQUENCE 100 AA; 11248 MW; D9973780A9B9DC23 CRC64;

Query Match 94.7%; Score 18; DB 1; Length 100;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSLF 5
 DB 39 SSLF 42

RESULT 9

YVYC_SCHPO STANDARD; PRT; 101 AA.

AC Q10219; 01-OCT-1996 (Rel. 34, last sequence update)

DT 01-OCT-1996 (Rel. 34, last sequence update)

DE 28-FEB-2003 (Rel. 41, last annotation update)

GN Very hypothetical protein C4H3.12c in chromosome I.

OS Schizosaccharomyces pombe (fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=972;

RX MEDLINE=21849401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,

RA Jones K., Jones L., Jones M., Leather S., McDonald S., Melean J.,

RA Oliver P., Neale S., Mungall K., Murphy L., Niblett D., Odell C.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

RA Woodward J., Voicikert G., Aert R., Robben J., Gymnoperez B.,

RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs W., Fritz C., Holzer E., Moestl D., Hilbert H.,

RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,

RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,

```

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mortier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerniti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurre P.,
RT "The genome sequence of Schizosaccharomyces pombe.",
RL Nature 415:871-880(2002).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z69380; CAA93351.1; -.
DR PIR; T38892; T38892.
DR GenDB_Spomb; SPAC4H3.12c; -.
KM Hypothetical protein; Transmembrane.
FT TRANSMEM 25 45 POTENTIAL.
SQ SEQUENCE 101 AA; 11914 MW; 6C686502AE2C0B09 CRC64;

Query Match 94.7%; Score 18; DB 1; Length 101;
Best Local Similarity 100.0%; Pred.No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
DB 96 SSLF 99

RESULT 10
YB67_MYCPN STANDARD; PRT; 101 AA.
ID YB67_MYCPN
AC P75316;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical lipoprotein MPM467 precursor (P01_orf101).
GN MPM467 OR MPM374.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
CX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfeich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.",
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -----
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE MG439 / MG440 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000036; AAB6022.1; -.
DR PIR; S73700; S73700.
DR InterPro; IPR001595; Lipoprotein_3.
DR Pfam; PF00938; Lipoprotein_3; 1.
DR ProDom; PD003276; Lipoprotein_3; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.

```

```

KM Hypothetical protein; Lipoprotein; Membrane; Signal;
KW Complete proteome.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 101 HYPOTHETICAL LIPOPROTEIN MPM467.
FT LIPID 26 26 N-ACYL DIHYDROIDE (POTENTIAL).
SQ SEQUENCE 101 AA; 11349 MW; 4DF6497B155C06 CRC64;

Query Match 94.7%; Score 18; DB 1; Length 101;
Best Local Similarity 100.0%; Pred.No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
DB 13 SSLF 16

RESULT 11
V078_FOWPV STANDARD; PRT; 103 AA.
ID V078_FOWPV
AC O9J5D2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein FPOV078.
GN FPOV078.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
CX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus.",
RL J. Virol. 74:3815-3831(2000).
CC -----
CC -1- SIMILARITY: BELONGS TO THE POXVIRUSES G3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF198100; AAF44422.1; -.
SQ SEQUENCE 103 AA; 12236 MW; BCE3BF93F3E79914 CRC64;

Query Match 94.7%; Score 18; DB 1; Length 103;
Best Local Similarity 100.0%; Pred.No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
DB 48 SSLF 51

RESULT 12
YJ02_YEAST STANDARD; PRT; 104 AA.
ID YJ02_YEAST
AC P47059;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 35, Last annotation update)
DE Hypothetical 12.4 kDa protein in DBP4-BE74 intergenic region.
GN YJL032W OR J1252.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
CX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Pohl T.M., Aljinovic G.;

```


Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; 249307; CAAB9322.1; -
DR PIR; S56804; S56804.
DR SGD; S0003568; YTL032W.
KW Hypothetical protein.
SQ SEQUENCE 104 AA; 12377 MW; 3A31474DF55E7C7B CRC64;

Query Match 94.7%; Score 18; DB 1; Length 104;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
DB 68 SSLF 71

RESULT 13

R44P CANMA

ID R44P CANMA STANDARD; PRT; 105 AA.

AC Q00477; Q00494; 01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE 60S ribosomal protein L44 P (L44) (L44 P-type).

GN L44P AND L44PIB

OS Candida maltosa (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; mitosporic Saccharomycetales; Candida.

ON NCBI_TaxID=5479;

RX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=IAM 12247;

RA Eshum M.;

RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

CC -1- MISCELLANEOUS: CYCLOHEXIMIDE SENSITIVE TYPE.

CC -1- SIMILARITY: BELONGS TO THE L44E FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; D43686; BAA0782.1; -

DR EMBL; D43687; BAA0784.1; -

DR InterPro; IPR000552; Ribosomal_L44E.

DR Pfam; PF00935; Ribosomal_L44; 1.

DR ProDom; PD002641; Ribosomal_L44E; 1.

DR PROSITE; PS01172; RIBOSOMAL_L44E; 1.

KW Ribosomal protein.

FT INIT MET

FT VARIANT 0

FT SEQUENCE 105 AA; 12037 MW; BD1368CE50B4938 CRC64;

SQ

Query Match 94.7%; Score 18; DB 1; Length 105;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5

DB 32 SSLF 35

RESULT 14

RL12 SULAC

ID RL12 SULAC STANDARD; PRT; 105 AA.

AC P08055; P35022; 01-AUG-1988 (Rel. 08, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE 50S ribosomal protein L12P ('A' type).

GN RPL12P.

OS Sulfolobus acidocaldarius.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Sulfolobus.

ON NCBI_TaxID=2285;

RX [1]

RP SEQUENCE.

RC STRAIN=ATCC 33909 / NCIB 11770 / DSM 639;

RA Matheson A.T., Louie K.A., Boeck A.;

RT "The complete amino acid sequence of the ribosomal A protein (L12)

RT from the archaeobacterium Sulfolobus acidocaldarius.",

RL FBS Letc. 231.331-335(1988).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=95055761; PubMed=7966335;

RA Ramirez C., Shimmun L.C., Leggett P., Matheson A.T.;

RT "Structure and transcription of the L11-L10-L12 ribosomal protein

RT gene operon from the extreme thermophilic archaeon Sulfolobus

RT acidocaldarius.",

RL Can. J. Microbiol. 244:242-249 (1994).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=9248688; PubMed=2497941;

RA Ramirez C., Shimmun L.C., Newton C.H., Matheson A.T., Dennis P.P.;

RT "Structure and evolution of the L11, L10, and L12 equivalent

RT ribosomal proteins in eubacteria, archaeobacteria, and eukaryotes.",

RL Can. J. Microbiol. 35:234-244 (1989).

RN [4]

RP ERRATUM.

RA Ramirez C., Shimmun L.C., Newton C.H., Matheson A.T., Dennis P.P.;

RL Can. J. Microbiol. 35:975-975 (1989).

CC -1- FUNCTION: SEEMS TO BE THE BINDING SITE FOR SEVERAL OF THE FACTORS

CC INVOLVED IN PROTEIN SYNTHESIS AND APPEARS TO BE ESSENTIAL FOR

CC ACCURATE TRANSLATION.

CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.

CC -1- CAUTION: WAS ORIGINALLY (REF.3) THOUGHT TO ORIGINATE FROM

CC S. SOLFARICUS STRAIN PI, BUT THE CULTURE WAS CONTAMINATED WITH

CC S. ACIDOCALDARIUS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X59038; CAA11765.1; -

DR PIR; S53651; RSUC12.

DR InterPro; IPR001813; 60S_ribosomal.

DR Pfam; PF00428; 60S_ribosomal; 1.

KW Ribosomal protein.

FT MOD RES 1

FT CONFLICT 46

FT SEQUENCE 105 AA; 11138 MW; 7B56E7CDB5A89C24 CRC64;

SQ

Query Match 94.7%; Score 18; DB 1; Length 105;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5

DB 101 SSLF 104

```

RESULT 15
PT17 STYPL STANDARD; PRT; 112 AA.
ID PT17 STYPL
AC P28209;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Protein-tyrosine phosphatase 17 (EC 3.1.3.48) (Fragment).
GN STY 17.
OS Styela plicata (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Unchordata; Ascidiacea;
OC Scolidobranchia; Styelidae; Styela.
OX NCBI_TaxID=7726;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9113172; PubMed=1704870;
RA Matthews R.J., Flores E., Thomas M.L.;
RT "Protein tyrosine phosphatase domains from the protochordate Styela
RL plicata.", Immunogenetics 33:33-41 (1991).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M38002; AAA2835.1; -
DR HSP; P18052; TYRO.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; TYR_PP.
DR Pfam; PF00102; Y_phosphatase.1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; PARTIAL.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; PARTIAL.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolyase.
FT NON_TER 1 1
FT NON_TER 112 112
SQ SEQUENCE 112 AA; 13081 MW; 2BE9DB8DCC5DB78 CRC64;

```

```

Query Match          94.7%; Score 18; DB 1; Length 112;
Best Local Similarity 100.0%; Pred.No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 SSUF 5
   ||||
Db 90 SSUF 93

```

Search completed: January 29, 2004, 14:53:12
 Job time : 4.23529 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OW protein - protein search, using sw model

Run on: January 29, 2004, 14:46:19 ; Search time 14.3265 Seconds
(without alignments)
86.441 Million cell updates/sec

Title: US-10-032-950-8

Perfect score: 19

Sequence: 1 XSSLF 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_orcanelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	94.7	17	4 Q95795	Q95795 homo sapien
2	18	94.7	27	2 Q97V1	Q97V1 straphylococ
3	18	94.7	27	12 Q8QSP5	Q8QSP5 hepatitis c
4	18	94.7	27	12 Q8QSM6	Q8QSM6 hepatitis c
5	18	94.7	27	12 Q9J5Y2	Q9J5Y2 hepatitis c
6	18	94.7	27	12 Q8QSM4	Q8QSM4 hepatitis c
7	18	94.7	27	12 Q9J5X8	Q9J5X8 hepatitis c
8	18	94.7	27	12 Q9J5X9	Q9J5X9 hepatitis c
9	18	94.7	27	12 Q9J5Y0	Q9J5Y0 hepatitis c
10	18	94.7	27	12 Q8QSM5	Q8QSM5 hepatitis c
11	18	94.7	27	12 Q9J5Y5	Q9J5Y5 hepatitis c
12	18	94.7	27	12 Q8QSM2	Q8QSM2 hepatitis c
13	18	94.7	27	12 Q9J5Y6	Q9J5Y6 hepatitis c
14	18	94.7	27	12 Q8QSM3	Q8QSM3 hepatitis c
15	18	94.7	27	12 Q9J5Y5	Q9J5Y5 hepatitis c
16	18	94.7	27	12 Q9J5X4	Q9J5X4 hepatitis c

17	18	94.7	27	12 Q9J5F4	Q9J5F4 hepatitis c
18	18	94.7	27	12 Q8QSM0	Q8QSM0 hepatitis c
19	18	94.7	27	12 Q8QSP4	Q8QSP4 hepatitis c
20	18	94.7	27	12 Q9J5X7	Q9J5X7 hepatitis c
21	18	94.7	27	12 Q8QSP7	Q8QSP7 hepatitis c
22	18	94.7	27	12 Q8QSM8	Q8QSM8 hepatitis c
23	18	94.7	27	12 Q9J5Y6	Q9J5Y6 hepatitis c
24	18	94.7	27	12 Q8QSM7	Q8QSM7 hepatitis c
25	18	94.7	27	12 Q8QSM1	Q8QSM1 hepatitis c
26	18	94.7	27	12 Q8QSL9	Q8QSL9 hepatitis c
27	18	94.7	34	8 Q62958	Q62958 picea abies
28	18	94.7	35	10 Q8QSM4	Q8QSM4 arabidopsis
29	18	94.7	35	11 Q9J5Y7	Q9J5Y7 mus musculus
30	18	94.7	35	12 Q9QJ70	Q9QJ70 tanapox vir
31	18	94.7	36	2 Q8VTS7	Q8VTS7 listeria in
32	18	94.7	36	2 Q8VTS5	Q8VTS5 listeria we
33	18	94.7	36	2 Q8VTR8	Q8VTR8 listeria iv
34	18	94.7	36	2 Q8VTS0	Q8VTS0 listeria mo
35	18	94.7	36	8 Q94RP7	Q94RP7 platyphora
36	18	94.7	39	3 Q8TRU9	Q8TRU9 saccharomyc
37	18	94.7	40	6 Q8TRU9	Q8TRU9 macropus ru
38	18	94.7	40	12 Q9J5Y0	Q9J5Y0 hepatitis c
39	18	94.7	40	12 Q9J5Y0	Q9J5Y0 hepatitis c
40	18	94.7	41	5 Q8IR66	Q8IR66 drosophila
41	18	94.7	41	8 Q8HE50	Q8HE50 ligyria simu
42	18	94.7	42	6 Q9J750	Q9J750 sus. scrofa
43	18	94.7	43	10 Q9SXL1	Q9SXL1 cucumis sat
44	18	94.7	43	16 Q8F826	Q8F826 leprospira
45	18	94.7	45	12 Q68616	Q68616 hepatitis c

ALIGNMENTS

RESULT 1

ID Q95795 PRELIMINARY; PRT; 17 AA.
AC Q95795;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE Advanced glycosylation end product-specific receptor (Fragment).
GN PAGE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Blazkova M., Kankova K.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065211; AAD15888.1; -.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 9 9
FT VARIANT 17 17
FT NON_TER 17 17
SQ SEQUENCE 17 AA; 1727 MW; 38CTEE8959E00D72 CRC64;

Query Match 94.7%; Score 18; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 2,1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSIF 5
Db 9 SSIF 12
RESULT 2
ID Q97V1 PRELIMINARY; PRT; 27 AA.
AC Q97V1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
 DE Glutamic acid-specific protease (EC 3.4.21.19) (Fragment).
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1280;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC12200;
 RX MEDLINE=93152142; PubMed=1369143;
 RA Kakedo S., Yoshikawa K., Tamaki M., Nakamura E., Teraoka H.;
 RT "Secretory expression of a glutamic-acid-specific endopeptidase
 (Ssepe) from Staphylococcus aureus ATCC12200 in Bacillus subtilis";
 RL Appl. Microbiol. Biotechnol. 38:226-233(1992).
 DR EMBL; D10369; BAA21889.1; -.
 KW Hydrolase; Protease.
 FT NON_TER 27
 SQ SEQUENCE 27 AA; 2769 MW; 707B6F95EC38BD4 CRC64;

Query Match 94.7%; Score 18; DB 2; Length 27;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
 ||||
 DB 9 SSLF 12

RESULT 3

Q8QSP5 PRELIMINARY; PRT; 27 AA.
 ID Q8QSP5;
 AC Q8QSP5;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
 DE E2 protein (genome polyprotein) (Fragment).
 GN E2.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=R3C4;
 RA Gaudy C., Moreau A., Veillon P., Lunel-Fabiani F., Goudreau A.;
 RT "Pretreatment analysis of HCV-1b HVR1 amino acid sequences to
 RT differentiate strains resistant and sensitive to antiviral therapy";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF473750; AAL86820.1; -.
 DR InterPro; IPR002531; HCV_NS1.
 DR Pfam; PF01560; HCV_NS1; 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 FT NON_TER 1
 FT NON_TER 27
 SQ SEQUENCE 27 AA; 2809 MW; 424DBFCD8ED5FBD7 CRC64;

Query Match 94.7%; Score 18; DB 12; Length 27;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
 ||||
 DB 17 SSLF 20

RESULT 4

Q8QSM6 PRELIMINARY; PRT; 27 AA.
 ID Q8QSM6;
 AC Q8QSM6;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
 DE E2 protein (genome polyprotein) (Fragment).

GN E2.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=R4C3;
 RA Gaudy C., Moreau A., Veillon P., Lunel-Fabiani F., Goudreau A.;
 RT "Pretreatment analysis of HCV-1b HVR1 amino acid sequences to
 RT differentiate strains resistant and sensitive to antiviral therapy";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF473759; AAL86839.1; -.
 DR InterPro; IPR002531; HCV_NS1.
 DR Pfam; PF01560; HCV_NS1; 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 FT NON_TER 1
 FT NON_TER 27
 SQ SEQUENCE 27 AA; 2674 MW; A7F417F44073C4D5 CRC64;

Query Match 94.7%; Score 18; DB 12; Length 27;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
 ||||
 DB 17 SSLF 20

RESULT 5

Q9J5Y2 PRELIMINARY; PRT; 27 AA.
 ID Q9J5Y2;
 AC Q9J5Y2;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20219428; PubMed=10756048;
 RA Manzin A., Solforosi L., Debiaggi M., Zara F., Tanzi E., Romano L.,
 RA Zanetti A.R., Clementi M.;
 RT "Dominant role of host selective pressure in driving hepatitis C virus
 RT evolution in perinatal infection";
 RL J. Virol. 74:4327-4334(2000).
 DR EMBL; AF192427; AAF65638.1; -.
 FT NON_TER 1
 FT NON_TER 27
 SQ SEQUENCE 27 AA; 2839 MW; 2C72910C79B20A04 CRC64;

Query Match 94.7%; Score 18; DB 12; Length 27;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
 ||||
 DB 17 SSLF 20

RESULT 6

Q8QSM4 PRELIMINARY; PRT; 27 AA.
 ID Q8QSM4;
 AC Q8QSM4;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
 DE E2 protein (genome polyprotein) (Fragment).
 GN E2.

OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R4C5;
 RA Gandy C., Moreau A., Veillon P., Lunel-Fabiani F., Goudeau A.;
 RT "Pretreatment analysis of HCV-1b HVR1 amino acid sequences to
 RT differentiate strains resistant and sensitive to antiviral therapy";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF473771; AAF6841.1; -
 DR InterPro: IP002531; HCV_NSI.
 DR Pfam: PF01560; HCV_NSI; I.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KM Polypeptide; Transmembrane.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2674 MW; A7F417F44073C4D5 CRC64;

Query Match 94.7%; Score 18; DB 12; Length 27;
 Best Local Similarity 100.0%; Pred.No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSLF 5
 Db 17 SSLF 20

RESULT 7
 ID 09J5X8 PRELIMINARY; PRT; 27 AA.
 AC 09J5X8
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Polypeptide (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20219428; PubMed=10756048;
 RA Manzini A., Solforosi L., Debiaggi M., Zara F., Tanzi E., Romano L.,
 RA Zanetti A.R., Clementi M.;
 RT "Dominant role of host selective pressure in driving hepatitis C virus
 RT evolution in perinatal infection";
 RL J. Virol. 74:4327-4334(2000).
 DR EMBL: AF192431; AAF65642.1; -
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2825 MW; 2C72839C79B20A04 CRC64;

Query Match 94.7%; Score 18; DB 12; Length 27;
 Best Local Similarity 100.0%; Pred.No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSLF 5
 Db 17 SSLF 20

RESULT 8
 ID 09J5X9 PRELIMINARY; PRT; 27 AA.
 AC 09J5X9
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Polypeptide (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20219428; PubMed=10756048;
 RA Manzini A., Solforosi L., Debiaggi M., Zara F., Tanzi E., Romano L.,
 RA Zanetti A.R., Clementi M.;
 RT "Dominant role of host selective pressure in driving hepatitis C virus
 RT evolution in perinatal infection";
 RL J. Virol. 74:4327-4334(2000).
 DR EMBL: AF192430; AAF65641.1; -
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2841 MW; 164299F29B20A1E CRC64;

Query Match 94.7%; Score 18; DB 12; Length 27;
 Best Local Similarity 100.0%; Pred.No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSLF 5
 Db 17 SSLF 20

RESULT 9
 ID 09J5Y0 PRELIMINARY; PRT; 27 AA.
 AC 09J5Y0
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Polypeptide (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20219428; PubMed=10756048;
 RA Manzini A., Solforosi L., Debiaggi M., Zara F., Tanzi E., Romano L.,
 RA Zanetti A.R., Clementi M.;
 RT "Dominant role of host selective pressure in driving hepatitis C virus
 RT evolution in perinatal infection";
 RL J. Virol. 74:4327-4334(2000).
 DR EMBL: AF192429; AAF65640.1; -
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2856 MW; 8C22814C79B20A02 CRC64;

Query Match 94.7%; Score 18; DB 12; Length 27;
 Best Local Similarity 100.0%; Pred.No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSLF 5
 Db 17 SSLF 20

RESULT 10
 ID 08Q5M5 PRELIMINARY; PRT; 27 AA.
 AC 08Q5M5
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE E2 protein (Genome polypeptide) (Fragment).
 GN E2.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.

STRAIN=R4C4;
 RC Gaudy C., Moreau A., Veillon P., Lunel-Fabiani F., Goudeau A.;
 RA "Pre-treatment analysis of HCV-1b HVR1 amino acid sequences to
 RT differentiate strains resistant and sensitive to antiviral therapy.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF473770; AAL86840.1; -.
 DR InterPro; IPR002531; HCV NS1.
 DR Pfam; PF01560; HCV NS1; 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KM Polypeptide; Transmembrane.
 FT NON_TER 1
 FT SEQUENCE 27 AA; 2674 MW; A7P417F44073C4D5 CRC64;
 SQ
 Query Match 94.7%; Score 18; DB 12; Length 27;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSUF 5
 ||||
 Db 17 SSUF 20

RESULT 11.
 Q91UF5 PRELIMINARY; PRT; 27 AA.
 ID Q91UF5
 AC Q91UF5; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Envelope protein (Genome polypeptide) (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Alberto S.-F.;
 RT "Influence of the dynamics of Hepatitis C virus quasispecies in the
 RT histological outcome of liver transplantation.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF221140; AAF7702.1; -.
 DR InterPro; IPR002531; HCV NS1.
 DR Pfam; PF01560; HCV NS1; 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KM Polypeptide; Transmembrane.
 FT NON_TER 1
 FT SEQUENCE 27 AA; 2715 MW; C8F6555E3CF59763 CRC64;
 SQ
 Query Match 94.7%; Score 18; DB 12; Length 27;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSUF 5
 ||||
 Db 16 SSUF 19

RESULT 12
 Q8QSM2 PRELIMINARY; PRT; 27 AA.
 ID Q8QSM2
 AC Q8QSM2; 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE E2 protein (Genome polypeptide) (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]

SEQUENCE FROM N.A.
 RC Gaudy C., Moreau A., Veillon P., Lunel-Fabiani F., Goudeau A.;
 RA "Pre-treatment analysis of HCV-1b HVR1 amino acid sequences to
 RT differentiate strains resistant and sensitive to antiviral therapy.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF473773; AAL86843.1; -.
 DR InterPro; IPR002531; HCV NS1.
 DR Pfam; PF01560; HCV NS1; 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KM Polypeptide; Transmembrane.
 FT NON_TER 1
 FT SEQUENCE 27 AA; 2674 MW; A7P417F44073C4D5 CRC64;
 SQ
 Query Match 94.7%; Score 18; DB 12; Length 27;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSUF 5
 ||||
 Db 17 SSUF 20

RESULT 13
 Q9Z068 PRELIMINARY; PRT; 27 AA.
 ID Q9Z068
 AC Q9Z068; 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE E2 region (Genome polypeptide) (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98411405; PubMed=9738045;
 RA Mizuno M., Higuchi T., Kamatsuse K., Esumi M.;
 RT "Genetic and serological evidence for multiple instances of
 RT unrecognized transmission of hepatitis C virus in hemodialysis
 RT units.";
 RL J. Clin. Microbiol. 36:2926-2931(1998).
 DR EMBL; AB001404; BAA35043.1; -.
 DR InterPro; IPR002531; HCV NS1.
 DR Pfam; PF01560; HCV NS1; 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KM Polypeptide; Transmembrane.
 FT NON_TER 1
 FT VARIANT 1
 FT NON_TER 1
 FT SEQUENCE 27 AA; 2749 MW; AB46C291E9E2FA44 CRC64;
 SQ
 Query Match 94.7%; Score 18; DB 12; Length 27;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSUF 5
 ||||
 Db 17 SSUF 20

RESULT 14
 Q8QSM3 PRELIMINARY; PRT; 27 AA.
 ID Q8QSM3
 AC Q8QSM3; 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE E2 protein (Genome polypeptide) (Fragment).
 OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R4C8;
 RA Gaudy C., Moreau A., Vellion P., Lunel-Fabiani F., Goudeau A.;
 RT "Pretreatment analysis of HCV-1b HVR1 amino acid sequences to
 RT differentiate strains resistant and sensitive to antiviral therapy."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF473772; AAF66842.1; -
 DR InterPro: IPR002531; HCV_NSI.
 DR Pfam: PF01560; HCV_NSI; 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2702 MW; A7F6C7F44073C4D5 CRC64;

Query Match 94.7%; Score 18; DB 12; Length 27;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSLF 5
 Db 17 SSLF 20

RESULT 15

O9J5Y5 PRELIMINARY; PRT; 27 AA.
 ID O9J5Y5
 AC O9J5Y5;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20219428; Pubmed=10756048;
 RA Manzini A., Solforosi L., Debiaggi M., Zara F., Tanzi E., Romano L.,
 RA Zanetti A.R., Clementi M.;
 RT "Dominant role of host selective pressure in driving hepatitis C virus
 RT evolution in perinatal infection."
 RL J. Virol. 74:4327-4334(2000).
 DR EMBL: AF192424; AAF65635.1; -
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2852 MW; 08B36BEC79B20A1E CRC64;

Query Match 94.7%; Score 18; DB 12; Length 27;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSLF 5
 Db 17 SSLF 20

Search completed: January 29, 2004, 14:56:52
 Job time : 15.9265 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:52:19 ; Search time 6.17647 Seconds
(without alignments)
34.252 Million cell updates/sec

Title: US-10-032-950-8

Perfect score: 19

Sequence: 1 XSLF 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*

2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*

3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*

4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*

5: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep:*

6: /cgn2_6/ptodata/1/1aa/backfilltest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	94.7	5	4	US-09-339-511-8
2	18	94.7	9	3	US-08-960-780-67
3	18	94.7	9	3	US-09-073-889-67
4	18	94.7	9	4	US-09-339-511-1
5	18	94.7	9	4	US-09-339-511-2
6	18	94.7	9	4	US-09-339-511-3
7	18	94.7	9	4	US-09-339-511-4
8	18	94.7	9	4	US-08-861-476C-3
9	18	94.7	12	1	US-08-156-552A-10
10	18	94.7	12	2	US-08-406-330-43
11	18	94.7	12	2	US-08-637-759B-201
12	18	94.7	12	2	US-08-556-597-43
13	18	94.7	12	3	US-08-871-355A-201
14	18	94.7	12	3	US-09-035-249A-10
15	18	94.7	12	4	US-09-201-945-201
16	18	94.7	14	1	US-08-156-552A-9
17	18	94.7	14	2	US-08-637-759B-153
18	18	94.7	14	3	US-08-871-355A-153
19	18	94.7	14	3	US-09-035-249A-9
20	18	94.7	14	4	US-09-201-945-153
21	18	94.7	15	2	US-08-553-257A-56
22	18	94.7	15	4	US-09-441-993-56
23	18	94.7	16	1	US-08-156-552A-8
24	18	94.7	16	3	US-09-035-249A-8
25	18	94.7	16	4	US-09-690-454-88
26	18	94.7	21	4	US-09-205-258-1062
27	18	94.7	22	1	US-08-468-824-7

28	18	94.7	25	4	US-09-149-476-369	Sequence 369, App
29	18	94.7	26	2	US-08-818-253-48	Sequence 48, Appl
30	18	94.7	26	3	US-09-094-359-16	Sequence 16, Appl
31	18	94.7	26	3	US-09-172-063-36	Sequence 36, Appl
32	18	94.7	26	3	US-08-818-252-48	Sequence 48, Appl
33	18	94.7	26	4	US-09-459-956-14	Sequence 14, Appl
34	18	94.7	26	4	US-09-316-019-21	Sequence 21, Appl
35	18	94.7	27	3	US-08-789-333F-24	Sequence 24, Appl
36	18	94.7	27	3	US-09-169-015-34	Sequence 34, Appl
37	18	94.7	27	3	US-09-133-944-33	Sequence 23, Appl
38	18	94.7	27	4	US-09-208-827-25	Sequence 25, Appl
39	18	94.7	27	4	US-08-787-739B-24	Sequence 24, Appl
40	18	94.7	27	4	US-09-157-748-26	Sequence 26, Appl
41	18	94.7	27	4	US-08-800-170-77	Sequence 77, Appl
42	18	94.7	27	4	US-09-626-581D-37	Sequence 37, Appl
43	18	94.7	27	4	US-09-415-765B-37	Sequence 37, Appl
44	18	94.7	27	4	US-09-626-580C-37	Sequence 37, Appl
45	18	94.7	27	6	5264477-1	Patent No. 5264477

ALIGNMENTS

```

RESULT 1
US-09-339-511-8
; Sequence 8, Application US/09339511
; Patent No. 6337385
; GENERAL INFORMATION:
; APPLICANT: Mult, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. 6337385ick, Richard P.
; APPLICANT: Beaver, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/09/339,511
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (1)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-09-339-511-8

Query Match          94.7%, Score 18; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. NO. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 SSLF 5
Db      2 SSLF 5

RESULT 2
US-08-960-780-67
; Sequence 67, Application US/08960780
; Patent No. 6204435
; GENERAL INFORMATION:
; APPLICANT: Peirceison, Jerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James

```



```

APPLICANT: Loewer, David
APPLICANT: Dullum, Charles Joseph
APPLICANT: Muller-Cohn, Judy
APPLICANT: Stamp, Lisa
TITLE OF INVENTION: No. 6204435a1 Pesticidal Toxins and Nucleotide
SEQUENCES: 134
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESS: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/960,780
FILING DATE: 30-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,848
FILING DATE: 30-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA-708
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-960-780-67

Query Match          94.7%; Score 18; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY      2 SSLF 5
DB      1 SSLF 4

RESULT 3
US-09-073-898-67
Sequence 67, Application US/09073898
Patent No. 6242669
GENERAL INFORMATION:
APPLICANT: Feltelson, Jerald S.
APPLICANT: Schepf, H. Ernest
APPLICANT: Narva, Kenneth E.
APPLICANT: Stockhoff, Brian A.
APPLICANT: Schmeltz, James
APPLICANT: Loewer, David
APPLICANT: Dullum, Charles Joseph
APPLICANT: Muller-Cohn, Judy
APPLICANT: Stamp, Lisa
APPLICANT: Morrill, George
APPLICANT: Finsrud-Jee, Stacey
TITLE OF INVENTION: No. 6242669a1 Pesticidal Toxins and Nucleotide
SEQUENCES: 144
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESS: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1

```

```

CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,898
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,848
FILING DATE: 30-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/960,780
FILING DATE: 30-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-708C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-073-898-67

Query Match          94.7%; Score 18; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY      2 SSLF 5
DB      1 SSLF 4

RESULT 4
US-09-339-511-1
Sequence 1, Application US/09339511
Patent No. 6337385
GENERAL INFORMATION:
APPLICANT: Muir, Tom
APPLICANT: Mayville, Patricia
APPLICANT: No. 6337385ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: Ji, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/09/339,511
CURRENT FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 1
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.

```

US-09-339-511-1

Query Match 94.7%; Score 18; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5

DB 6 SSUF 9

RESULT 5
 US-09-339-511-2
 ; Sequence 2, Application US/09339511
 ; Patent No. 6337385
 ; GENERAL INFORMATION:
 ; APPLICANT: Muir, Tom
 ; APPLICANT: Mayville, Patricia
 ; APPLICANT: No. 6337385ick, Richard P.
 ; APPLICANT: Beavis, Ronald
 ; APPLICANT: Ji, Gangyong
 ; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
 ; FILE REFERENCE: 600-1-231N
 ; CURRENT APPLICATION NUMBER: US/09/339,511
 ; CURRENT FILING DATE: 1999-06-24
 ; PRIOR APPLICATION NUMBER: 60/090,402
 ; PRIOR FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; NAME/KEY: VARIANT
 ; LOCATION: (5)
 ; OTHER INFORMATION: Xaa represents any amino acid at this position.
 ; US-09-339-511-2

Query Match 94.7%; Score 18; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5

DB 6 SSUF 9

RESULT 6
 US-09-339-511-3
 ; Sequence 3, Application US/09339511
 ; Patent No. 6337385
 ; GENERAL INFORMATION:
 ; APPLICANT: Muir, Tom
 ; APPLICANT: Mayville, Patricia
 ; APPLICANT: No. 6337385ick, Richard P.
 ; APPLICANT: Beavis, Ronald
 ; APPLICANT: Ji, Gangyong
 ; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
 ; FILE REFERENCE: 600-1-231N
 ; CURRENT APPLICATION NUMBER: US/09/339,511
 ; CURRENT FILING DATE: 1999-06-24
 ; PRIOR APPLICATION NUMBER: 60/090,402
 ; PRIOR FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 9
 ; TYPE: PRT

Query Match 94.7%; Score 18; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

OTHER INFORMATION: peptide

NAME/KEY: VARIANT

LOCATION: (5)

OTHER INFORMATION: Xaa represents any amino acid at this position.

US-09-339-511-3

Query Match 94.7%; Score 18; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5

DB 6 SSUF 9

RESULT 7
 US-09-339-511-4
 ; Sequence 4, Application US/09339511
 ; Patent No. 6337385
 ; GENERAL INFORMATION:
 ; APPLICANT: Muir, Tom
 ; APPLICANT: Mayville, Patricia
 ; APPLICANT: No. 6337385ick, Richard P.
 ; APPLICANT: Beavis, Ronald
 ; APPLICANT: Ji, Gangyong
 ; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
 ; FILE REFERENCE: 600-1-231N
 ; CURRENT APPLICATION NUMBER: US/09/339,511
 ; CURRENT FILING DATE: 1999-06-24
 ; PRIOR APPLICATION NUMBER: 60/090,402
 ; PRIOR FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; NAME/KEY: VARIANT
 ; LOCATION: (5)
 ; OTHER INFORMATION: Xaa represents any amino acid at this position.
 ; US-09-339-511-4

Query Match 94.7%; Score 18; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5

DB 6 SSUF 9

RESULT 8
 US-08-861-476C-3
 ; Sequence 3, Application US/08861476C
 ; Patent No. 6447786
 ; GENERAL INFORMATION:
 ; APPLICANT: New York University Medical Center
 ; TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUEREUS
 ; FILE REFERENCE: 63753/7
 ; CURRENT APPLICATION NUMBER: US/08/861,476C
 ; CURRENT FILING DATE: 1997-05-22
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Version 3.0
 ; SEQ ID NO 3
 ; LENGTH: 9
 ; TYPE: PRT

Query Match 94.7%; Score 18; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORGANISM: Staphylococcus aureus
US-08-861-476C-3

Query Match 94.7%; Score 18; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5
|||
Db 6 SSUF 9

RESULT 9

US-08-156-552A-10
Sequence 10, Application US/08156552A
Patent No. 572615

GENERAL INFORMATION:

APPLICANT: Bokoch, Gary M

ATTORNEY/AGENT INFORMATION:

TITLE OF INVENTION: REGULATION OF OXIDATIVE BURST USING

TITLE OF INVENTION: LKMG-DERIVED PEPTIDES AND ANALOGS

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: The Scripps Research Institute, Office of

STREET: 10666 No. 572615th Torrey Pines Road, TPC 8

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/156,552A

CLASSIFICATION: 514

FILING DATE: 15-NOV-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/102,944

FILING DATE: 02-AUG-1993

ATTORNEY/AGENT INFORMATION:

NAME: Logan, April C.

REGISTRATION NUMBER: 33,950

REFERENCE/DOCKET NUMBER: SCRF 281.1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-554-2937

TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-156-552A-10

Query Match 94.7%; Score 18; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5
|||
Db 1 SSUF 4

RESULT 10

US-08-406-330-43
Sequence 43, Application US/08406330

Patent No. 5817748

GENERAL INFORMATION:

APPLICANT: Miller, Jonathan L.

ATTORNEY/AGENT INFORMATION:

NAME: Vacki, A.

TITLE OF INVENTION: MINOTOPES AND ANTI-MINOTOPES OF
TITLE OF INVENTION: HUMAN PLATELET GLYCOPROTEIN Ib/IX
NUMBER OF SEQUENCES: 81

CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon, Hargrave, Devans & Doyle

STREET: Clinton Square, P.O. Box 1051

CITY: Rochester

STATE: New York

COUNTRY: USA

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/406,330

FILING DATE:

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Timian, Susan J.

REGISTRATION NUMBER: 34,103

REFERENCE/DOCKET NUMBER: 20884/100

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 263-1636

TELEFAX: (716) 263-1600

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-406-330-43

Query Match 94.7%; Score 18; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5
|||
Db 3 SSUF 6

RESULT 11

US-08-637-759B-201
Sequence 201, Application US/08637759B

Patent No. 5876931

GENERAL INFORMATION:

APPLICANT: David William Holden

TITLE OF INVENTION: Identification of Genes

NUMBER OF SEQUENCES: 501

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patrea L. Pabst

STREET: 2800 One Atlantic Center

STREET: 1201 West Peachtree Street

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30309-3450

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/637,759B

FILING DATE: 03-MAY-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB95/02875

FILING DATE: 11-DEC-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPS 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 201:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-637-7599-201

Query Match 94.7%; Score 18; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
Db 8 SSLF 11

RESULT 12
US-08-556-597-43
Sequence 43, Application US/08556597
Patent No. 5877155
GENERAL INFORMATION:
APPLICANT: Miller, Jonathan L.
TITLE OF INVENTION: NIMTOPES AND ANTI-NIMTOPES OF
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/556,597
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/406,330
FILING DATE: 17-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Timian, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 20884/101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1636
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-556-597-43

Query Match 94.7%; Score 18; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 66;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
Db 3 SSLF 6

RESULT 13
US-08-871-355A-201
Sequence 201, Application US/08871355A
Patent No. 6015669
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871,355A
FILING DATE: 09-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPS 101 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 201:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-871-355A-201

Query Match 94.7%; Score 18; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
Db 8 SSLF 11

RESULT 14
US-09-035-248A-10
Sequence 10, Application US/09035248A
Patent No. 6184203
GENERAL INFORMATION:
APPLICANT: Bokoch, Gary M
TITLE OF INVENTION: REGULATION OF OXIDATIVE BURST USING
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:

ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSER: Patent Counsel
STREET: 10666 No. 6184203th Torrey Pines Road, TPC 8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/035,249A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/156,552
FILING DATE: 15-NOV-1993
APPLICATION NUMBER: US 08/102,944
FILING DATE: 02-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Logan, April C.
REGISTRATION NUMBER: 33,950
REFERENCE/DOCKET NUMBER: SCRF 281.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-035-249A-10

Query Match 94.7%; Score 18; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 SSLF 5
|||
Db 1 SSLF 4

RESULT 15
US-09-201-945-201
Sequence 201, Application US/09201945
Patent No. 6342215
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSER: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/201,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/637,759
FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPKS 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 201:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-201-945-201

Query Match 94.7%; Score 18; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 SSLF 5
|||
Db 8 SSLF 11

Search completed: January 29, 2004, 15:03:23
Job time : 7.17647 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 14:47:39 ; Search time 14.1912 Seconds
(without alignments)
73.223 Million cell updates/sec

Title: US-10-032-950-8

Perfect score: 19

Sequence: 1 XSLIF 5

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	94.7	5	13	US-10-032-950-8
2	18	94.7	7	12	US-10-022-066-211
3	18	94.7	9	10	US-09-850-351A-67
4	18	94.7	9	12	US-10-022-066-266
5	18	94.7	9	13	US-10-032-950-1
6	18	94.7	9	13	US-10-032-950-2
7	18	94.7	9	13	US-10-032-950-3
8	18	94.7	9	13	US-10-032-950-4
9	18	94.7	9	15	US-10-201-444-3
10	18	94.7	10	12	US-09-572-270A-470
11	18	94.7	11	12	US-10-293-086-12
12	18	94.7	11	15	US-10-097-175-86
13	18	94.7	12	12	US-10-435-666-1
14	18	94.7	13	12	US-10-300-694A-5
15	18	94.7	13	15	US-10-174-613-31

16	18	94.7	14	12	US-10-158-825-59	Sequence 59, App1
17	18	94.7	14	15	US-10-158-847-59	Sequence 59, App1
18	18	94.7	15	9	US-09-759-010-10	Sequence 10, App1
19	18	94.7	16	12	US-10-062-599-88	Sequence 88, App1
20	18	94.7	16	15	US-10-062-831-88	Sequence 88, App1
21	18	94.7	16	15	US-10-225-567A-1613	Sequence 1613, App
22	18	94.7	18	9	US-09-864-761-40860	Sequence 40860, A
23	18	94.7	18	11	US-09-866-066-28	Sequence 28, App1
24	18	94.7	19	9	US-09-864-761-31256	Sequence 3756, A
25	18	94.7	19	9	US-09-864-761-41274	Sequence 41274, A
26	18	94.7	19	11	US-09-776-124A-221	Sequence 221, App
27	18	94.7	19	12	US-10-160-162-339	Sequence 339, App
28	18	94.7	19	12	US-09-820-649-339	Sequence 339, App
29	18	94.7	20	12	US-09-962-756-742	Sequence 742, App
30	18	94.7	20	12	US-09-962-756-744	Sequence 744, App
31	18	94.7	20	12	US-09-962-756-746	Sequence 746, App
32	18	94.7	20	12	US-09-962-756-748	Sequence 748, App
33	18	94.7	20	12	US-10-253-471-742	Sequence 742, App
34	18	94.7	20	12	US-10-253-471-744	Sequence 744, App
35	18	94.7	20	12	US-10-253-471-746	Sequence 746, App
36	18	94.7	20	12	US-10-253-471-748	Sequence 748, App
37	18	94.7	21	12	US-09-933-767-1062	Sequence 1062, App
38	18	94.7	21	12	US-09-962-756-602	Sequence 602, App
39	18	94.7	21	12	US-10-253-471-602	Sequence 602, App
40	18	94.7	21	15	US-10-023-282-1062	Sequence 1062, App
41	18	94.7	24	12	US-09-833-245-377	Sequence 377, App
42	18	94.7	25	11	US-09-809-391-369	Sequence 369, App
43	18	94.7	25	12	US-09-882-171-369	Sequence 369, App
44	18	94.7	25	12	US-10-029-386-32388	Sequence 32388, A
45	18	94.7	26	10	US-09-999-745-21	Sequence 21, App1

ALIGNMENTS

RESULT 1
US-10-032-950-8
Sequence 8, Application US/10032950
Publication No. US2002007453A1
GENERAL INFORMATION:
APPLICANT: Muir, Tom
APPLICANT: Mayville, Patricia
APPLICANT: No. US2002007453A1ick, Richard P.
APPLICANT: Beavis, Ronald
APPLICANT: Ji, Guangyong
TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
TITLE OF INVENTION: INTERFERENCE
FILE REFERENCE: 600-1-231N
CURRENT APPLICATION NUMBER: US/10/032,950
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 60/090,402
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
NAME/KEY: VARIANT
LOCATION: (1)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-8

Query Match 94.7% Score 18; DB 13; Length 5;
Best Local Similarity 100.0%; Pred. No. 6.9e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
DB 2 XSLIF 5

```

RESULT 2
US-10-022-066-231
; Sequence 231, Application US/10022066
; Publication No. US20030166057A1
; GENERAL INFORMATION:
; APPLICANT: HILDEBRAND, WILLIAM H.
; APPLICANT: PRILLIMAN, KILEY RAE
; TITLE OF INVENTION: METHOD AND APPARATUS FOR THE PRODUCTION OF ANTIGENS AND
; FILE REFERENCE: 6680.034
; CURRENT APPLICATION NUMBER: US/10/022,066
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/256,410
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/256,409
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 09/465,321
; PRIOR FILING DATE: 1999-12-17
; PRIOR APPLICATION NUMBER: 09/974,366
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 231
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-022-066-231

Query Match          94.7%; Score 18; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.9e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
   ||||
Db 4 SSLF 7

RESULT 3
US-09-850-351A-67
; Sequence 67, Application US/09850351A
; Patent No. US2002010080A1
; GENERAL INFORMATION:
; APPLICANT: Feltelson, Jerald S.
; Schneft, H. Ernest
; Narva, Kenneth E.
; Stockhoff, Brian A.
; Schmeltz, James
; Loewer, David
; Dullum, Charles Joseph
; Muller-Cohn, Judy
; Stamp, Lisa
; Morrill, George
; TITLE OF INVENTION: No. US2002010080A1el Pesticidal Toxins and Nucleotide
; Sequences Which Encode These Toxins
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/09/850,351A
; FILING DATE: 07-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/073,898
; FILING DATE: 06-May-1998
; APPLICATION NUMBER: US 08/960,780
; FILING DATE: 30-Oct-1997
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-Oct-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA-708CD1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 67:
US-09-850-351A-67

Query Match          94.7%; Score 18; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.9e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
   ||||
Db 1 SSLF 4

RESULT 4
US-10-022-066-266
; Sequence 266, Application US/10022066
; Publication No. US20030166057A1
; GENERAL INFORMATION:
; APPLICANT: HILDEBRAND, WILLIAM H.
; APPLICANT: PRILLIMAN, KILEY RAE
; TITLE OF INVENTION: METHOD AND APPARATUS FOR THE PRODUCTION OF ANTIGENS AND
; FILE REFERENCE: 6680.034
; CURRENT APPLICATION NUMBER: US/10/022,066
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/256,410
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/256,409
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 09/465,321
; PRIOR FILING DATE: 1999-12-17
; PRIOR APPLICATION NUMBER: 09/974,366
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 266
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-022-066-266

Query Match          94.7%; Score 18; DB 12; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.9e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
   ||||

```

Db 6 SSUF 9

RESULT 5

US-10-032-950-1
; Sequence 1, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mult, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-1

Query Match 94.7%; Score 18; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.9e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5

Db 6 SSUF 9

RESULT 6

US-10-032-950-2
; Sequence 2, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mult, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-2

Query Match 94.7%; Score 18; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.9e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5

Db 6 SSUF 9

RESULT 7

US-10-032-950-3
; Sequence 3, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mult, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: VARIANT
; LOCATION: (5)
; OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-3

Query Match 94.7%; Score 18; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.9e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSUF 5

Db 6 SSUF 9

RESULT 8

US-10-032-950-4
; Sequence 4, Application US/10032950
; Publication No. US20020077453A1
; GENERAL INFORMATION:
; APPLICANT: Mult, Tom
; APPLICANT: Mayville, Patricia
; APPLICANT: No. US20020077453A1ick, Richard P.
; APPLICANT: Beavis, Ronald
; APPLICANT: Ji, Guangyong
; TITLE OF INVENTION: NOVEL STAPHYLOCOCCUS PEPTIDES FOR BACTERIAL
; FILE REFERENCE: 600-1-231N
; CURRENT APPLICATION NUMBER: US/10/032,950
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/090,402
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-10-032-950-4

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Peptide
NAME/KEY: VARIANT
LOCATION: (5)
OTHER INFORMATION: Xaa represents any amino acid at this position.
US-10-032-950-4

Query Match 94.7%; Score 18; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.9e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
Db 6 SSLF 9

RESULT 9
US-10-201-444-3
Sequence 3, Application US/10201444
Publication No. US20030078378A1
GENERAL INFORMATION:
APPLICANT: New York University Medical Center
TITLE OF INVENTION: BLOCKING EXPRESSION OF VIRULENCE FACTORS IN S. AUREUS
FILE REFERENCE: 63753/7
CURRENT APPLICATION NUMBER: US/10/201,444
CURRENT FILING DATE: 2002-07-23
PRIOR APPLICATION NUMBER: US/08/861,476
PRIOR FILING DATE: 1997-05-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 9
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-201-444-3

Query Match 94.7%; Score 18; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.9e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
Db 6 SSLF 9

RESULT 10
US-09-572-270A-470
Sequence 470, Application US/09572270A
Publication No. US20030148368A1
GENERAL INFORMATION:
APPLICANT: Proteom Ltd
TITLE OF INVENTION: Inter- complementary peptide listing
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/572,270A
CURRENT FILING DATE: 2000-05-17
NUMBER OF SEQ ID NOS: 1144
SOFTWARE: ProPatent version 1.0
SEQ ID NO 470
LENGTH: 10
TYPE: PRT
ORGANISM: Arabidopsis Thaliana
OTHER INFORMATION: Sequence located in Unknown at 10-19 and may interact with
US-09-572-270A-470

Query Match 94.7%; Score 18; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
Db 7 SSLF 10

RESULT 11
US-10-293-086-12
Sequence 12, Application US/10293086
Publication No. US20030134310A1
GENERAL INFORMATION:
APPLICANT: Cytoc, Thomas P.
TITLE OF INVENTION: Cellular Kinase Targets and Inhibitors.
FILE REFERENCE: 50036/048002
CURRENT APPLICATION NUMBER: US/10/293,086
CURRENT FILING DATE: 2003-03-25
PRIOR APPLICATION NUMBER: US 60/337,990
PRIOR FILING DATE: 2001-11-13
NUMBER OF SEQ ID NOS: 144
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 11
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Phosphorylation consensus sequence
US-10-293-086-12

Query Match 94.7%; Score 18; DB 12; Length 11;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
Db 7 SSLF 10

RESULT 12
US-10-097-175-86
Sequence 86, Application US/10097175
Publication No. US20030045680A1
GENERAL INFORMATION:
APPLICANT: JOYAL, JOHN L.
APPLICANT: MUELLER, JOHN
APPLICANT: OZA, VISHA B.
APPLICANT: FINDERS, MARK A.
TITLE OF INVENTION: PEPTIDIC MODULATORS OF THE ANDROGEN RECEPTOR
FILE REFERENCE: PEI-110
CURRENT APPLICATION NUMBER: US/10/097,175
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/275,240
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/352,399
PRIOR FILING DATE: 2002-01-28
NUMBER OF SEQ ID NOS: 102
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 86
LENGTH: 11
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Androgen Receptor Binding Polypeptides
US-10-097-175-86

Query Match 94.7%; Score 18; DB 15; Length 11;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSLF 5
Db 2 SSLF 5

RESULT 13
US-10-435-666-1
Sequence 1, Application US/10435666
Publication No. US2004006001A1

```

; GENERAL INFORMATION:
; APPLICANT: CARTER, Daniel C.
; TITLE OF INVENTION: FERRITIN FISSION PROTEINS FOR USE IN VACCINES AND OTHER APPLICATIO
; FILE REFERENCE: P07624WMOO/BAS
; CURRENT APPLICATION NUMBER: US/10/435,666
; PRIOR FILING DATE: 2003-05-12
; PRIOR APPLICATION NUMBER: 60/379,145
; PRIOR FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-435-666-1

Query Match          94.7%; Score 18; DB 12; Length 12;
Best Local Similarity 100.0%; Pred. No. 2,4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 SSLF 5
      |||
      3 SSLF 6

Db

RESULT 14
US-10-300-694A-5
; Sequence 5, Application US/10300694A
; Publication No. US20030185870A1
; GENERAL INFORMATION:
; APPLICANT: Duke University
; APPLICANT: Grinstaff, Mark W.
; APPLICANT: Kenan, Daniel J.
; APPLICANT: Walsh, Elisabeth B.
; APPLICANT: Middleton, Crystan
; TITLE OF INVENTION: INTERFACIAL BIOMATERIALS
; FILE REFERENCE: 180/143/2
; CURRENT APPLICATION NUMBER: US/10/300,694A
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: US 60/331,843
; PRIOR FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Polystyrene-binding peptide 5
US-10-300-694A-5

Query Match          94.7%; Score 18; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 2,4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 SSLF 5
      |||
      2 SSLF 5

Db

RESULT 15
US-10-174-613-31
; Sequence 31, Application US/10174613
; Publication No. US20030114354A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL POTASSIUM CHANNEL WITH HOMOLOGY T
; TITLE OF INVENTION: Ether-a-go-go FAMILY, HEAG2
; FILE REFERENCE: D0123 NP
; CURRENT APPLICATION NUMBER: US/10/174,613
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,378
; PRIOR FILING DATE: 2001-06-19

```

```

; PRIOR APPLICATION NUMBER: US 60/300,614
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 31
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-174-613-31

Query Match          94.7%; Score 18; DB 15; Length 13;
Best Local Similarity 100.0%; Pred. No. 2,4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 SSLF 5
      |||
      1 SSLF 4

Db

Search completed: January 29, 2004, 15:01:47
Job time : 15.1912 secs

```